

CONFORMATION OF PROTEINS

BY SERGE N. TIMASHEFF¹ AND MARINA J. GORBUNOFF¹

*Eastern Regional Research Laboratory,
Utilization Research and Development Division,
Agricultural Research Service, U.S. Department of Agriculture,
Philadelphia, Pennsylvania*

When contemplating the question of the recent advances in the field of the conformations of proteins, one is overwhelmed by the vastness of the subject and the surging flood of new knowledge. The questions and answers which are constantly generating a vast literature are: (a) What are the conformations of various proteins? (b) Why do the proteins assume such conformations? (c) How do the conformations change? (d) How do we detect these conformations and their changes? While in a limited review it would be a hopeless task to undertake a comprehensive survey of events that have occurred even most recently, the approach to this general problem seems to be following several lines: (a) the direct determination by X-ray diffraction of the three-dimensional structure of proteins in the crystalline state; (b) calculations of possible structures from data on bond distances and angles, potential barriers to free rotation and side-chain interactions; (c) the determination by various techniques of the secondary and tertiary structures of proteins in solution; and (d) the theoretical examination and experimental probing of the thermodynamic factors which determine the conformations assumed by proteins in aqueous solution and the changes which occur in them with changes in environment.

The first topic encompasses the determination of the three-dimensional space coordinates of proteins in the crystal state. The second includes the calculations, pioneered by Ramachandran, of the structures which are permitted for given sequences of amino acids. These have been the subjects of excellent reviews recently (1, 2) and will not be discussed here. Over the past ten years, the elucidation of the secondary and tertiary structures of proteins in solution has been the subject of much effort, principally using spectroscopic techniques. Quite recently, significant advances have been made in the frequency ranges both of the ultraviolet and of the infrared (principally in the range of the amide I and amide II bands); these will be discussed in turn. Another challenging question which seems worthy of review is concerned with the degree of burial within the molecule, or of exposure to solvent, of chromophoric side-chains; much progress has been made on this subject using the techniques of differential ultraviolet absorption spectroscopy and accessibility of the groups to protons and to various

¹ Presently at the Graduate Department of Biochemistry, Brandeis University, Waltham, Massachusetts.

modifying reagents. Finally, the degree of success attained recently in predicting changes in protein conformation with changes in environment from a consideration of the energies of transfer of amino acid residues and peptide bonds from water to various nonaqueous or highly perturbed aqueous media deserves a detailed discussion; this topic will not be reviewed here, however, because of space limitations and because its discussion would require a complete chapter. This review, therefore, will be restricted to some aspects of the experimental determination of the secondary and tertiary structures of proteins.

ULTRAVIOLET OPTICAL ROTATION

Although this subject has been reviewed quite thoroughly in recent years (1-4) and was the topic of a chapter in these reviews a year ago (2), sufficient progress has been made in the past year to warrant a summary and analysis. In particular, important advances have been made in the optical rotatory dispersion (ORD) characterization of the pleated-sheet, or β structure, the circular dichroism (CD) of polypeptides in various conformations, and the critical treatment of ORD and CD results.

ORD of the β conformations.—Attempts at the determination of the Moffitt & Yang parameters of the various β conformations go back to the pioneering studies of Fasman & Blout (5), and Imahori (6), who, in 1960, published results on poly-O-acetyl-L-serine in dichloroacetic acid (DCA)-chloroform solutions, in which this polypeptide undergoes a random $\rightarrow\beta$ transition as the amount of DCA is decreased. From infrared absorption flow dichroism experiments, Imahori & Yahara (7) identified this structure as a cross- β , i.e., an antiparallel chain (APC) in which the interchain hydrogen bonds are running parallel to the axis of fiber orientation. The b_0 value for this material was found to be zero, while the a_0 parameter was large and positive (5-7), tending toward +700 (7). Bradbury, Elliott & Hanby (8), working with poly-O-benzyl-L-serine, reported $a_0 = +600$, $b_0 = +190$, for what was again identified as a cross- β structure in a DCA-chloroform mixture. More recently, Ikeda, Maeda & Isemura (9) have examined poly-S-carbobenzoxymethyl-L-cysteine in the same solvent mixture, reporting a_0 and b_0 values of +400 and 0 for an antiparallel-chain (APC) β structure. Similar values of a_0 and b_0 were confirmed for some of these polypeptide derivatives by Anufrieva et al. (10). Most recently, a theoretical examination of this structure by Volkenstein & Zubkov (11) in terms of the Moffitt, Fitts & Kirkwood (12) theory of optical rotation has established that an APC β structure should have a b_0 value of close to zero. It appears, thus, that in examining the conformation of a protein or polypeptide in terms of the Moffitt & Yang (13) or Shechter & Blout (14) two-term equations, the intrinsic value of b_0 (also $A_{193} - A_{225}$) for the antiparallel-chain β conformation may be taken to be zero as a close approximation. An unequivocal choice of a_0 is more difficult. From the reported values, this quantity seems to have a large positive value (probably

between 400 and 700). This parameter, however, is known to be highly sensitive to interactions with solvent and side-chain effects (15, 16) and must be approached with a great deal of caution. The conclusion, however, that b_0 of the APC- β conformation is close to zero renders this parameter, as well as $A_{193} - A_{225}$, more useful for estimating the contents of α -helix since b_0 is rather insensitive to solvent and side-chain effects (15, 16) and for an unordered (or random) conformation, it assumes small values. It should be borne in mind, however, that no similar data are available for the parallel-chain β -structure, which has been predicted by theoretical calculations (17) to have a different rotational spectrum in the far ultraviolet region. Similar studies in the far ultraviolet (below 260 m μ) are highly complicated by the fact that most solvents used with the above polypeptide derivatives are not transparent in this spectral region. The first ORD spectrum of a β structure in the region of absorption of peptide bonds was reported by Blout & Shechter (18) for a film of poly-L-isoleucine; this pattern is characterized by a peak at 206 m μ and a trough at 230 m μ . A similar spectrum was reported by Davidson, Tooney & Fasman (19) on a film of poly-L-serine, who found a peak at 203 m μ and a trough at 233 m μ .

It is known that at alkaline pH (11, 12), heated poly-L-lysine undergoes an α -helix $\rightarrow \beta$ transition. This fact was exploited by Davidson, Tooney & Fasman (19) and Sarkar & Doty (20) who reported almost identical ORD patterns for the polylysine β structure, with a trough at 230 m μ , a peak at 205 m μ , and residue rotations of -6,300 and 23,000 to 29,000 respectively. Iizuka & Yang (21) have found that *Bombyx mori* L. silk fibroin in the β conformation has a peak at 205 m μ and troughs at 230 m μ and 190 m μ . In both cases, the β structure has been shown by infrared spectroscopy to be of the antiparallel chain type (20-22).

The observation that an APC β structure ORD spectrum has a trough at 230 to 233 m μ renders quite dangerous (2, 19, 23, 24) the frequent practice of using this trough as a diagnostic criterion for the presence of α -helix. Furthermore, a very weak trough at 232 to 235 m μ can be seen as well in the ORD spectrum of polypeptides in random conformation (23, 25). The positive peak at lower wavelengths appears to be a better conformational criterion, since its positions in the α -helical, APC β and random (or unordered) conformations are at 199, 206, and ca. 190 m μ , respectively (23, 25, 26).

Circular dichroism.—The method of circular dichroism (CD), which is closely related to ORD, possesses the advantage of having relatively narrow positive and negative bands which can be resolved with greater ease than the corresponding infinitely broad ORD bands. Since the ORD Cotton effects can be calculated band by band from the CD spectrum with the Kronig-Kramers transform (27-30), comparison of the two techniques can be made quite readily. The presently available values of the positions, ellipticities, $[\theta]$, and rotational strengths, R^k , of the pertinent bands are summarized in Table I. CD spectra of polypeptides in the α -helical con-

TABLE I^a
CIRCULAR DICHROISM OF POLYPEPTIDES IN VARIOUS CONFORMATIONS

Conformation	Band Position	$[\theta'] \times 10^{-3}$	$R_L \times 10^{-40}$
	$m\mu$	$\text{deg-cm}^2/\text{decimole}$	$\text{erg-cm}^2\text{-rad}$
α -Helix	221-2	-30.4	-(17-22)
	207	-28-29	-(12-14)
	190-2	52-55	30-40
Unordered	235-8	~ -0.2	-(0.05-0.15)
	217	~ 2	1-2
	196-7	-(25-35)	-(14-16)
APC- β (in H ₂ O) (in SDS or Methanol)	217	-(14-17)	-11
	195	21	14
	217	-(6-9)	-7
	197	22-30	12-24
	190	16	4
Parallel- β	216	-	-(4-7)
	181	-	21-24
	236	-4 ^b	$\sim -2^b$
Poly-L-proline I in <i>n</i> -propanol	214-6	58 ^b	28 ^b
	200-3	-30 ^b	-18 ^b
	221	13-16	5-7
Poly-L-proline II in H ₂ O	207	-(35-40)	-(32-33)
	223	2	~ 1
Collagen	198	-22	-13
	188	+	

^a In this and subsequent tables no references are given to the sources of material cited; these are all identified in the text.

^b Values uncorrected for refractive index.

formation have been reported over the past five years (20, 22, 30-35) and compared successfully with theoretical predictions and the corresponding ORD data. This aspect of CD has been reviewed in 1966 (2). During the past year, much progress has been made in the similar characterization of the APC β structure, using poly-L-lysine in aqueous media (20, 22) and silk fibroin (21), while, quite recently, a theoretical analysis has become available (17). In all three experimental studies (20-22), a negative band has been found at 217 $m\mu$, while in two (21, 22), a strong positive band has been reported at 196 $m\mu$. Although in the α -helix only one of three bands (at 206 $m\mu$) is sensitive to the nature of the solvent, in the APC β structure the intensity of both bands appears to be a strong function of the polarity of the medium. In silk fibroin (21), the 196 $m\mu$ band increased in intensity and that at 217 decreased when the solvent was changed from

50 per cent to 93 per cent methanol. A similar decrease in the 217 m μ peak (20, 36) and increase in the 196 m μ peak (36) was observed for poly-L-lysine, when the solvent was 0.06 to 0.12 *M* sodium dodecyl sulfate (SDS). It is interesting to add, that similar experiments with poly-D-lysine have given spectra identical in amplitude but opposite in sign to those obtained with poly-L-lysine (55).

Pysh (17) has carried out a theoretical analysis of the ultraviolet optical properties of polypeptides in both the antiparallel and parallel chain β conformations. In the case of the APC β structure, the major π - π^* transition band is located at 195 m μ with a weaker band at 198 m μ ; the n - π^* transition results in weak negative rotation. These calculations are in agreement with experimental observations (20-22, 36); furthermore, it is interesting to note that in 0.12 *M* SDS, the low wavelength positive CD band splits (36), as predicted by theory. The parallel-chain β -structure should have a spectrum different from that of the APC β -structure (17): its π - π^* Cotton effect is expected at 181 m μ with a weak negative n - π^* Cotton effect near 216 m μ . Thus, CD, as well as infrared spectroscopy (see below), can serve as a method for differentiating between the two pleated sheet configurations.

The CD spectrum of polypeptides in the unordered form has been described in several publications (32, 34-36). It is characterized by a strong negative band at 196 m μ , a weak positive band at 217 m μ and very weak negative absorption centered at about 238 m μ . The electron transitions associated with the first two bands have been discussed (2, 34), although the weak negative absorption between 230 and 245 m μ is difficult to interpret (23) and may be only the trailing limb of the strong negative band at 196 m μ (37). The observed CD spectrum of poly-L-lysine (36) in the random conformation is in good agreement with that calculated (23) for poly-L-glutamic acid.

Three other polypeptide conformations are the two types of poly-L-proline helix and the collagen triple helix. Their ORD spectra (23, 36, 38-41) are characterized by a peak at 223 m μ and a trough at 208 m μ for polyproline I and a peak and trough at 192 and 217 m μ for polyproline II; collagen has a deep trough at 206 m μ and essentially invariant weak negative rotation between 230 and 260 m μ (38). The CD spectra of the three conformations have been reported (23, 36). Polyproline I (36) has a strong positive absorption, maximal at 214 m μ , strong negative absorption, maximal at 198 m μ , and weak negative absorption, centered at 236 m μ . The CD spectrum of polyproline II consists of a strong negative absorption peaking at 205 m μ (36) and weak positive absorption maximal 230 m μ (23, 36); this spectrum can be resolved into two bands at 207 and 221 m μ (23, 36). The CD spectrum of calf skin collagen has a strong negative band at 198 m μ and a weak positive absorption at 223 m μ (36). In all three cases, the calculated ORD curves are in good agreement with experi-

mental data. Pysh (42) has carried out a theoretical examination of these spectra, predicting for polyproline I two π - π^* CD bands of equal intensity, a positive one at 216 m μ , and a negative one at 204 m μ ; and two well-separated π - π^* CD bands for poly-L-proline II, one negative at 201 m μ and a weaker positive one at 215 m μ . For the collagen triple helix, Pysh predicts two π - π^* CD bands, a negative one at 198 m μ and a weaker positive one near 190 m μ . In all three cases, the n - π^* contribution is very small. Comparison of the theoretically calculated spectra with the experimental ones indicates that, except for the collagen 190 m μ band, all have been observed (36).

Conformational analysis.—The determination of the CD spectra of the principal structures assumed by polypeptides has rendered possible attempts at an analysis of the conformations of proteins and polypeptides in various media. Such structural analysis can be complicated seriously by the effects of solvent, side-chain interactions and aggregation. Tomimatsu, Vitello & Gaffield (43) have examined the depth of the 233 m μ trough of poly-L-glutamic acid in α -helical conformation as a function of polypeptide aggregation. Their finding that with aggregation the amplitude of the 233 m μ trough increases, explains, at least in part, earlier reports of varying depths of the 233 m μ trough in poly-L-glutamic acid (44, 45). They interpret this change in terms of the proposal by Schuster (46) who associated the increase in levorotation with aggregation to an increase in local dielectric constant upon extensive lateral overlapping and close association of helical chains. In the case of globular proteins, a change in a_0 , but no b_0 , has been observed during the aggregation of β -lactoglobulin (47), in which a_0 becomes less negative; this has been attributed to the transfer of additional amino acid residues to the hydrophobic interior of the protein (47) or, alternately, to the formation of additional APC β structure (48).

Carver, Shechter & Blout (16, 23) have carried out a systematic analysis of the factors involved in the interpretation of the ORD and CD data of proteins. First, using a modified Moscowitz equation, they calculated from ORD data the CD spectra of the α -helical, random and poly-L-proline II conformations. Then, they analyzed the conditions necessary for interpreting ORD data in terms of protein conformations. In order for such an analysis to be meaningful, four conditions must be fulfilled: (a) the backbone peptide bonds are the only source of optical activity contributing to the rotatory parameter; (b) the rotatory parameters must be independent of the number of residues in a segment of a given conformation; (c) the rotatory parameters must be insensitive to effects of solvents and side-chain interactions; and (d) the experimental error in the determination of the parameter in question must be negligible relative to the maximum contributions of the various structures. An examination of various methods of conformational analysis, namely, the single wavelength method, various dispersion methods, and the comparison of complete dispersion curves, led

to the conclusions that while the last method is found to be the one of choice, α -helix contents can be estimated quite reliably from the b_0 and ($A_{193} - A_{225}$) parameters, which are linearly related. Thus, while at present an exact determination of a protein conformation from ORD is impossible, it does appear that relative amounts of various conformations may be estimated from a very careful analysis of ORD and CD data.

Examining the same problem, Fasman and co-workers (49) have calculated quite recently the ORD curves to be expected in the far ultraviolet from various combinations of α -helical, APC β , and unordered conformations. Using their data on poly-L-lysine as standards for the various structures, they have obtained a family of curves showing peaks, troughs and cross-over points at wavelengths between 190 and 240 $m\mu$. When examining these curves, one is struck by the fact that they encompass almost all the types of ORD spectra that have been reported for proteins. In the light of present-day knowledge, this type of empirical approach to an estimate of the conformational composition of a protein may be the most realistic.

Another attempt at deducing the conformations of proteins from ORD data has been made by Troitskii (50) who, using equations similar to those of Urnes & Doty (3), constructed a set of nomograms for determining the conformational composition of proteins from values of a_0 and b_0 (or related parameters). Some 43 proteins have been examined in this way. Since the ORD data used, however, were not corrected for contributions of the chromophoric bands between 260 and 300 $m\mu$ and, furthermore, since some of the intrinsic values of a_0 and b_0 used for the three conformations are outside of the most likely range, their results must be regarded with a great deal of caution.

Using a similar approach to those of Fasman (49) and Carver & Colleagues (16, 23), a conformational analysis of β -lactoglobulin has been carried out from combined ORD and CD data (36, 48). First, the visible range wavelength rotation was corrected for the calculated contribution of chromophoric residue dichroic bands between 270 and 310 $m\mu$. This correction had practically no effect on a_0 but resulted in a 20 per cent change in b_0 . Using these corrected values of the Moffitt & Yang parameters, the most likely contents of α -helical, APC β and unordered conformations were calculated and the CD and ORD curves between 185 and 250 $m\mu$ were resolved, in terms of standard reference curves for polypeptides, into contributions from these three structures, with the conclusion that this protein contains a small amount (5-15 per cent) of α -helix, about half unordered structure and the rest APC β conformation. It must be pointed out, however, that in the case of β -lactoglobulin, it was already known from independent infrared absorption experiments (51) that this protein contains a considerable amount of APC β structure.

Other proteins that, up to the writing of this review, had been examined by CD in the region below 240 $m\mu$, include insulin (36, 52), myosin (53),

lysozyme (20, 36), γ -globulin (20), deoxyribonuclease (20), α_1 -acid glycoprotein (20), bovine serum albumin (36), α_s -casein (36), bovine carbonic anhydrase (36), phosvitin (36, 54), carboxypeptidase A (55), soybean trypsin inhibitor (55), ribonuclease (55), and chymotrypsin (55). The spectra of these display various degrees of complexity and thus lend themselves only to tentative qualitative interpretation. No detailed structural analysis, however, has been attempted on these proteins as yet.

In conclusion, the present state of knowledge of ORD and CD has put into proper perspective the serious difficulties which are inherent in the application of these techniques to protein conformation. While a beginning at qualitative conformational analysis seems possible, using the outlined approaches (16, 23, 36, 48), it is quite evident that much more information is necessary on the ORD and CD spectra of standard polypeptides in known conformations in various media. In particular, it would seem desirable to have detailed and highly precise ORD and CD spectra of those globular proteins, the structures of which are known from X-ray diffraction analysis.

INFRARED SPECTROSCOPY

The infrared absorption spectra of polypeptides and fibrous proteins (in particular the bands related to the C=O and -N-H groups) have been the object of studies for more than twenty years, resulting in a considerable amount of empirical correlation between conformations and band positions. Of the many bands associated with the peptide bonds, the amide I (1600 to 1700 cm^{-1} , C=O stretching) and amide II (1500 to 1550 cm^{-1} , N-H in-plane bending, C-N stretching) bands are the most amenable to conformational analysis and have been used for diagnostic purposes, since their positions and polarizations are strong functions of the polypeptide conformation. About five years ago, an important advance was made when Miyazawa (56) carried out a theoretical analysis of the positions and polarizations of these bands, in terms of intramolecular interactions which occur in different conformations, and made corresponding band assignments for the α -helical, parallel-chain (PC) pleated sheet, antiparallel-chain (APC) pleated sheet and unordered (or random) conformations (56-58). This work was the subject of a very lucid review three years ago (1) and will not be discussed in detail here.

In 1962, Miyazawa's calculations were taken up by Krimm (59), who modified certain band assignments and extended the calculation to the parallel-chain polar sheet and to polyglycine II. In Table II, the calculated amide I frequencies are compared with those observed for various conformations under a number of conditions. Krimm's modifications include the choice of $\nu_0 = 1520 \text{ cm}^{-1}$ [instead of 1535 (57)] for the unperturbed frequency of the amide II band, and the introduction of an additional interaction term into the calculation of the amide I and amide II frequencies for the α -helix. Applying these results to a number of fibrous proteins, Krimm

TABLE II
PRINCIPAL AMIDE I FREQUENCIES FOR VARIOUS CONFORMATIONS, CM^{-1}

Conformation	Mode	Calculated	Globular Proteins in			Polypeptides in	
			Solid	H ₂ O	D ₂ O	H ₂ O	D ₂ O
Unordered	ν_0	1658		1656	1643		1643, 1645
APC- β	$\nu_{ } (0, \pi)$	1685	1690	1690	1675	1690	1680
	$\nu_{ } (\pi, 0)$	1632	1630	1632	1632	1616	1611
PC- β	$\nu_{ } (0, 0)$	1648					
	$\nu_{\perp} (\pi, 0)$	1632					
Polar sheet	$\nu (0, 0)$	1648					
α -Helix	$\nu_{ } (0)$	1650	1652	1652	1650		1635, 1640

has concluded that the most likely conformation of feather keratin is a polar chain, i.e., a chain in which successive peptide groups have essentially similar spatial orientations, being related to each other by a glide plane, while in the nonpolar chain (the usual β structure) such groups are related by a twofold screw axis. From an analysis of the infrared spectrum of polyglycine II, Krimm concluded that the three chains are predominantly parallel to each other, since the calculated and observed amide II band positions would disagree strongly with the antiparallel-chain arrangement, which had been left as a definite possibility (60). Analysis of the infrared spectrum of oriented TMV particles (61) led to the conclusion that TMV protein contains no extended chain (β) structures, but is a mixture of α -helical and unordered conformations, a result in qualitative agreement with conclusions drawn from ORD (62). Furthermore, the α -helices in the protein are inclined at an angle of 60° with respect to the axis of the virus particle.

As an example of analyses which may be expected in the future, one should cite Tsuboi's very careful application of infrared dichroism to the conformation of the α -form of poly- γ -benzyl-L-glutamate (63). From an analysis of the spectrum of oriented films, between 500 and $3,800\text{ cm}^{-1}$, a number of band assignments both of the main-chain and of side-chains have been made and the orientations of peptide, ester and phenyl groups with respect to the helical axis have been determined. This analysis has led Tsuboi to conclude that in the poly- γ -benzyl-L-glutamate film the amide group is rotated 10° in its plane with respect to the normal bond directions in an α -helix; this makes the C-N bonds more nearly horizontal and produces a significant change in pitch of the α -helix. This result seems of particular significance in view of the recent detailed analyses of the helical portions of myoglobin and lysozyme which display many departures from an exact α -helical structure. In a later paper, Tsuboi (64) discussed some problems involved in the analysis of polypeptide spectra: he pointed out

that poly- γ -benzyl-L-glutamate gives more than 50 well defined absorption bands in the 400 to 4000 cm^{-1} region, of which only a small fraction have been analyzed in detail. After discussing the amide B band of polyglycine (3088 cm^{-1}) and amino wagging bands, Tsuboi examined bands related to (side chain) (side chain) interactions. To do this, he compared the spectra given by films cast from poly- γ -benzyl-L-glutamate, its optical antipode, 50 to 50 mixtures of the two and the copolymer, as well as similar films of the corresponding methyl esters; any spectral difference observed can be attributed entirely to the difference of the interaction between α -helices of the same sense and that between the α -helices of opposite sense. In the case of the benzyl ester, the pronounced dichroism of phenyl bands indicates that these groups are in fixed orientation in the DL blend, whereas much less regularity exists in the L-ester. Furthermore, the shift to higher frequency in the positions of several bands indicates stronger interchain interactions of $\text{CH}_2\cdots\text{CH}_2$ groups in the DL blend of the benzyl ester and of H \cdots H groups in a similar blend of the methyl ester.

With the recent emergence of the β structure as an important component of proteins, detailed analysis of the various bands associated with such conformations is highly desirable. Thus, the recent investigation of polyglycine by Suzuki et al. (65) and the assignment of bands for form I by Fukushima, Ideguchi & Miyazawa (66) is of great importance. Polyglycine I has the form of an APC pleated sheet (67); polyglycine II exists as a triple helix (60). Suzuki et al. have measured very carefully the infrared spectra of polyglycine I and polyglycine II, as well as the N-deuterated, C-deuterated, N, S-deuterated and N¹⁵-substituted polymers in the two forms, and, on the basis of these spectra, have obtained a nearly complete set of assignments of the bands. Fukushima et al. have calculated the normal vibrations of polyglycine I and its N-deuterated derivative in order to elucidate the nature of such vibrations characteristic of the antiparallel zig-zag conformation. This almost complete analysis of the spectra of polyglycine I gives an understanding of the absorption bands given by the antiparallel β chains free from any contribution of side-chains. It should facilitate, therefore, the analysis of similar spectra to be obtained in the future from other polypeptides in similar conformation; one can contemplate the systematic introduction of various side-chains and a progressive analysis of their contributions to the infrared spectrum.

The data of Suzuki et al. permit some direct comparisons of band shifts to be made with changes of conformation (form I \rightarrow II) and with N-deuteration (the latter is quite important for studies on proteins in aqueous medium, as will be shown below). The conformational change is accompanied by a shift of all the major bands and, thus, in principle, should be amenable to study by infrared spectroscopy. The effect of deuteration on the amide I and II bands of polyglycine I is: amide I (1685 \rightarrow 1680; 1636 \rightarrow 1629); amide II (1517 \rightarrow 1475); for polyglycine II: amide I (1644 \rightarrow 1639); amide

II (1554 \rightarrow 1476). These band shifts are quite significant and must be kept in mind when examining the spectra of proteins in D₂O solution. While yielding important basic information for the analysis of protein spectra, the results of Suzuki et al. are also an excellent example of the complexity of the general problem. One must remember that proteins contain a number of different conformations, each one making its specific contribution to the total spectrum. As an example, one might consider the complex spectrum that would be given by an equimolar mixture of polyglycines I and II.

A number of papers have appeared on the characterization of the "cross- β structure" (68), in which the chains run perpendicular to the fiber axis or direction of film orientation. This results in opposite dichroism of the bands to what is observed within the usual β structures (69). In 1960, Bradbury et al. (70) reported that some films cast from poly- β -*n*-propyl-L-aspartate, poly- β -benzyl-L-aspartate, and poly- γ -benzyl-L-glutamate of low molecular weight assume the cross- β structure. Since then a considerable amount of work has been done on poly-L-serine derivatives. Fasman & Blout (5) found that poly-O-acetyl-L-serine films exist in the antiparallel β structure. Imahori & Yahara (7, 71) reported that, depending on the degree of stroking, a film of this polymer cast from trifluoroacetic acid could assume either a "cross- β ", or a normal β structure (throughout their papers, these authors refer to "parallel β " when, quite evidently, they mean "noncross- β ", since the presence of a band at 1690 cm⁻¹ fixes the assignment of this structure as antiparallel β). The presence of cross- β conformation was confirmed by X-ray diffraction analysis (72). Infrared spectra of this polymer in a 25 per cent dichloroacetic acid (DCA) to 75 per cent chloroform solution (5, 7) indicate the presence of β structure, which appears to be of the cross- β nature as shown by infrared flow dichroism experiments (7). This result is particularly important, since it can be used in the correlation of other solution properties with a known cross- β conformation. Bradbury, Elliott & Hanby (8) confirmed the existence of cross- β structure in poly-O-benzyl-L-serine both in film and chloroform-DCA solution. Ikeda, Maeda & Isemura (9) have found that the dichroism of infrared bands of oriented poly-S-carbobenzoxymethyl-L-cysteine films corresponds to a cross- β structure, while in chloroform solution with less than 10 per cent DCA, the spectrum changed progressively from that of a random coil (amide I varying from 1655 to 1658 cm⁻¹ as DCA decreases from 4 to 0.5 per cent) to a β form (amide I varying from 1636 to 1629 cm⁻¹ as DCA decreases from 4 to 0.5 per cent). These authors discuss the band shifts in the transition region in terms of the gradual formation of hydrogen bonds as the solvent becomes poorer. As has been mentioned above, this system, like some of the previous ones, was used to characterize the ORD properties of the APC β structure. The most definitive study of a cross- β structure is that of Fraser et al. (73) who compared highly crystalline oriented films of sequential poly-L-alanylglycine with Bombyx mori silk fibroin. From the posi-

tions and orientations of the bands, the synthetic polymer was found to be in the cross- β structure while the protein had a normal APC β structure. Comparison of the two spectra demonstrates also the effect of side-chains on the spectrum of a protein.

With the vast advances made during the last few years in the assignment of absorption bands to various conformations, it would appear reasonable, at present, to attempt a qualitative structural examination of globular proteins. This is a task of great difficulty since, (a) most proteins contain a number of different conformations, which give rise to infrared bands that are diffuse and overlapping. (See Table II); these conformations interact with each other as seen with poly-S-carbobenzoxy-L-cysteine. (b) Globular proteins are handled best in aqueous medium; water, however, has a strong absorption band at 1650 cm^{-1} , i.e., directly in the amide I region. Therefore, solution studies must be carried out in D_2O which is transparent in this range; this, however, introduces band shifts attributable to N-deuteration, as seen with polyglycine. Drying out of the proteins in films may result in denaturation. Polarization measurements are essentially impossible, since stretching of a protein film would, most probably, result in some structural changes. A possible approach to such studies involves the suspension of the native crystalline proteins in some inert material; even this, however, does not preclude some denaturation, the result of interaction between the protein and the medium [it is known that high concentrations of ethylene glycol, for example, induce changes in protein conformation (74)]. A saving factor might be the fact that in crystals only the molecules in the surface layer come into contact with the medium, while those in the interior might be protected from possible denaturation. Attempts at such structural analysis go back to the early studies of Elliott and co-workers (69, 75-77) who examined several proteins either as dry films or in paste form. They described several $\alpha \rightarrow \beta$ transformations, a particularly interesting result being their discovery that the insulin fibrils exist in the form of a cross- β structure. In 1959 Beer et al. (78) surveyed a number of proteins, pointing out the complications which are inherent in such a structural analysis. Elliott, Hanby & Malcolm (79) examined lysozyme in cast films and found a band at 1660 cm^{-1} , which led them to conclude that the protein in this form had an unordered structure. Hamaguchi (80) published a spectrum of lysozyme in D_2O (at unspecified conditions), showing an amide I band maximal at 1650 cm^{-1} , with the hint of a shoulder at 1630 cm^{-1} , suggesting the presence of an α -helical structure with some β structure contribution. Shigorin & Zubov (81) have published the spectra (between $2,700$ and $3,600\text{ cm}^{-1}$) of the "globular" and "fibrous" forms of casein, with the conclusion that the former is α -helical while the latter has a β structure. Noticing the analogy between the optical rotatory dispersion properties of γ -globulin with those of poly-O-acetyl-L-serine, Imahori (82) suggested that this protein must have a cross- β structure. In an elegant experiment,

he prepared γ -globulin fibrils by precipitating human γ -globulin with horse antibody to this protein. An oriented film of these fibrils displayed an infrared dichroic spectrum typical of a cross APC β structure, with a shoulder indicative of some α -helical or unordered conformation. The spectrum obtained with unpolarized light gave a broad amide I band at 1650 cm^{-1} explaining why Imahori & Momoi (83) had not seen the β band in earlier experiments.

Recently, the amide I bands of several proteins were examined in D_2O solution by Timasheff & Susi (51). It was found that native myoglobin has a sharp band at 1650 cm^{-1} , characteristic for α -helical structure, as should be expected; native β -lactoglobulin (at pD 7.5), as well as a film of this protein, has a strong band at 1632 cm^{-1} with a shoulder at 1685 cm^{-1} , indicating the presence of considerable APC β structure; the identification by infrared spectroscopy of this β structure in β -lactoglobulin made possible the subsequent interpretation of its ORD (48) and CD (36, 84) spectra (see above). The unordered proteins α_s -casein and denatured β -lactoglobulin at pD 11.5 were found to have amide I bands at 1643 cm^{-1} , considerably shifted from the Miyazawa & Krimm assignment (Table II). Following these results, a systematic study was initiated on the comparison of the amide I band positions of these proteins in H_2O and D_2O solutions and in films, and also of the effects of denaturation on these positions (85, 86). It was found that in the case of the globular, tightly folded proteins, myoglobin and β -lactoglobulin, the spectra were essentially identical whether taken in film, H_2O , D_2O or in a Nujol suspension of the crystalline protein. The unordered proteins, α_s -casein and alkali-denatured β -lactoglobulin, had maxima in H_2O at 1655 cm^{-1} in agreement with the assignment of Table II. The band positions in films were very close to those found in H_2O in all cases. A similar examination of poly-L-lysine in the three conformations in H_2O and D_2O solutions revealed that deuteration shifts the bands toward lower frequencies, as had been reported for N-deuterated polyglycine (65). An essentially similar situation exists with polyglutamic acid (85). Analysis of the frequencies observed in H_2O and D_2O , in terms of the approach of Miyazawa (56) and Krimm (59), has shown that the pattern of D_2O -induced band shifts in open structures and its absence in tight ones is completely reasonable. The screening of a number of proteins in D_2O solution and in Nujol suspensions (36, 86) indicates that infrared spectroscopy can be quite useful in conformational analysis and, in particular, in following conformational changes. In most cases, however, the amide I bands are not sharply definitive for any single conformation, and it is in a comparison with the results of other techniques (such as circular dichroism, which by themselves again do not give sufficiently clear-cut answers) that infrared spectroscopy is most effective. Of the proteins examined, however, a few can be cited where the amide I band indicates the predominance of a given structure. Thus, bovine serum albumin is predominantly helical (1649 cm^{-1})

as is lysozyme (1651 cm^{-1}); native insulin probably has much helix (1654 cm^{-1}) in a constrained state; native bovine carbonic anhydrase contains much β structure (1636 cm^{-1}), while the acid denatured protein is more helical (1647 cm^{-1}). Methanolic denaturation of β -lactoglobulin confirms the $\beta \rightarrow \alpha$ transformation (shift from 1632 to 1648 cm^{-1}), while alkaline denaturation leads to disordering of the structure (shift to 1643 cm^{-1} in D_2O and 1656 cm^{-1} in H_2O).

The progress made over the last five years in the assignment of bands in the spectra of polypeptides has rendered possible the systematic use of infrared spectroscopy in the study of protein conformations. It would appear that, in tightly folded globular proteins, deuteration does not cause any band shifts. In structures where most residues come in contact with solvent, such as disordered proteins or structured and unstructured polypeptides, the band positions are highly sensitive to the nature of the medium, underlining the difficulty inherent in an extrapolation of band positions obtained from polypeptides in D_2O to globular proteins. Furthermore, the existence of a variety of β conformations which can be distinguished most easily by polarization experiments, renders this task even more difficult. The presence of a band in the 1685 to 1700 cm^{-1} region, however, can be used diagnostically for the identification of an antiparallel-chain β -conformation (59).

BURIED AND EXPOSED GROUPS

With the recent development of methods for the determination of the amino acid sequences of proteins and the three-dimensional structure of proteins in the crystalline state, the detailed knowledge of the exact positions in space of various side-chains has been greatly advanced. The question, however, of the location, in the three-dimensional network, of specific amino acid residues in a protein molecule in solution still remains to be solved (87). That such groups and their relative positions in space are of major importance to biological activity has been demonstrated in a number of studies. For example, Singer (88) has established the interaction of a carboxyl with an amino group in certain antigen-antibody reactions, while a number of studies have led to the implication of two histidines and one lysine residue in the active site of ribonuclease (89). The problem of the location of active residues in a protein molecule has been attacked by a number of techniques, with chemical modification and difference spectroscopy being among the most fruitful.

In the three-dimensional structure of proteins, individual amino acid residues may be present either on the surface of the molecule and accessible to the solvent medium, or buried in the internal hydrophobic regions. Such burial may be the result either of the folding of the polypeptide chain in which the residues of interest are incorporated or of intermolecular aggregation, in which groups present in the surfaces of intermolecular contacts

are removed from the solvent. Naturally, a rigorous classification into completely exposed and completely buried residues is a great oversimplification, since many residues belong to an intermediate class, neither fully exposed nor fully buried.

As a result of their location within the three-dimensional structure of a protein, groups are characterized by varying degrees of reactivity toward specific reagents. Groups on the surface, i.e., normal groups, resemble simple peptides in their reactivity, while those buried in the interior have reactivities lower than those of simple peptides and often are not reactive at all. Groups in intermediate classes display a wide spectrum of reactivity which probably reflects their degree of exposure to the solvent medium and the effects of vicinal residues on a given reaction. Thus, since the secondary and tertiary structures determine the behavior of amino acid residues incorporated into a protein molecule, it becomes possible to probe the environment of such groups by studying their reactivity with specific reagents.

This review comprises a discussion of some selective modification studies of only three amino acids, namely histidine, tryptophan, and tyrosine, along with a survey of solvent perturbation spectroscopy studies. Wherever possible, an attempt has been made to evaluate the reported methods in order to provide some basis for judging the environmental sensitivity of the discussed reagents. The conclusions which can be drawn at present on the basis of these comparisons perforce must be considered as tentative, since, in most cases, there is insufficient overlap between results obtained with the various methods.

Histidine.—Although it has been known for many years that the diazotization of proteins gives colored products, it is only recently, since Horinishi et al. (90) introduced diazo-1-H-tetrazole (DHT) as a histidine modifying reagent, capable of discriminating between free and buried histidine residues, that this reaction has been applied successfully to probing the surface of proteins.

The experimental procedure is based on the spectrophotometric determination of the amount of histidinebisazo-1-H-tetrazole ($\lambda_{\text{max}} = 480 \text{ m}\mu$) which is formed when a protein solution is treated with diazo-1-H-tetrazole at pH 8.8 and 16° C. The optical density at $\lambda_{\text{max}} = 480 \text{ m}\mu$ is plotted as a function of the diazo compound concentration. The highest critical concentration for any given protein is determined by that concentration above which a shoulder becomes recognizable at 550 $\text{m}\mu$. This shoulder is attributable to tyrosinebisazo-1-H-tetrazole which begins to form once the exposed histidine residues have been transformed into the bisazo form. The buried histidine residues do not react because of their inaccessibility. Plateaus in plots of the absorptivity of 480 $\text{m}\mu$ vs. the azo compound concentration indicate discrete levels of histidine residue reactivity. Where a plateau is poorly defined, the absorptivity at the critical concentration is used.

The states of the histidine residues in bacitracin (90), insulin (90),

TABLE III
THE STATE OF HISTIDINE RESIDUES IN PROTEINS

Protein	Titration		Reaction with DHT			Photo-oxidation
	pK	Number	Free	Buried	Denat. protein	
Bacitracin			1			
Lysozyme	6.8	1	1			1
Insulin	6.4	2	2			2
BSA	6.9	18	18		18	
Trypsin	normal		3		3	
Trypsinogen			2+1		3	
Chymotrypsinogen	6.7	2	(1)	(1)	1 (4 hr); 2 (18 hr)	
α -Chymotrypsin			1+1		2	1
α -CT-BGME			1+1			
DIP-chymotrypsin			1	(1)		
Cytochrome c: Fe ⁺⁺⁺ (horse heart)	6.8	1+2 abnormal	1+1+1		3	1 fast + 1 slow
Fe ⁺⁺			1+1+1			
Cytochrome c: Fe ⁺⁺⁺ (yeast)			2+1+1			
Fe ⁺⁺			2+1+1			
Ribonuclease	6.5	4	1+1+1+1		4	2+1
Ribonuclease+AMP or GMP			1+1+1+1			
Ribonuclease+CMP			1+2+1			
Ribonuclease+UMP			2+1+1			
Human serum β_1 -lipoprotein			2-3	(16-18)		
Human serum β_1 -lipoprotein (delipidated)			2-3	(16-18)		
Soybean trypsin inhibitor			2			

lysozyme (90), bovine serum albumin (BSA) (90), trypsin (90), trypsinogen (90), chymotrypsinogen (90), α -chymotrypsin, α -chymotrypsin in the presence of substrate (91), DIP-chymotrypsin (91), cytochrome c (92), ribonuclease (93) and carboxypeptidase A (93a) have been investigated using DHT, and those in human serum β_1 -lipoprotein (94) by diazotization with diazobenzene arsenate using a procedure which goes back to that of Tabachnick & Sobotka (95). The results of these studies, together with supporting data from the literature, are compiled in Table III.

As can be seen from Table III, all the histidine residues in bacitracin, lysozyme, insulin, BSA, and trypsin react with diazo-1-H-tetrazole and, thus, are normal. These results are in agreement with titration data on lysozyme, insulin, and BSA, which indicate that all the histidine residues are indeed normal in these proteins (96). Furthermore, it has been shown that one histidine residue in lysozyme (97) and two histidine residues in insulin (97) are destroyed by photooxidation.

Trypsinogen has two types of histidine residues. After alkaline denaturation, however, only one type is found. This transformation indicates that the difference in reactivity of the histidine residues in the native protein is

related to the secondary and tertiary structures of the molecule; the histidine residues become alike once their environmental differences have been destroyed. Comparison of trypsin with its zymogen leads to the conclusion that the activation of the zymogen is accompanied by the release to contact with solvent of one partly buried histidine residue and loosening of the structure in its neighborhood. This is supported further by the fact that the reactivities of the histidine residues in native and denatured trypsin are almost identical while, in trypsinogen, the histidine residues become much more reactive after denaturation, i.e., the reaction is completed at a much lower concentration of the diazo compound. It is known that the histidines of trypsin (98) titrate between pH 5 and 7 and, therefore, are normal although the reported reactivity of only a single residue with the specific reagent 1-chloro-3-tosyl-7-amino-2-heptanone (99) might point to some difference between the environments of the three residues.

A situation similar to that of the trypsinogen-trypsin pair is found in the case of α -chymotrypsin, its zymogen, its substrate complex with benzoylglycine methyl ester (BGME), and DIP-chymotrypsin (90, 91). In chymotrypsinogen (90), the concentration curve shows no plateau. There is some leveling off at about 0.65 histidine residues and a value of 1.2 residues is obtained from the critical concentration. After 4 hr of denaturation at pH 13, a plateau is obtained at the level of one residue, while an 18 hr. exposure raised the plateau to the level of two reactive residues. Two types of histidines are found in α -chymotrypsin, the difference disappearing after denaturation at pH 13. The reactivity of the histidine residues does not seem to be affected by the introduction of benzoylglycine methyl ester (BGME) substrate (91) which does not change the reactivity curves. In DIP-chymotrypsin only one residue is accessible to the azo reagent (91); the second one only reacts to the extent of about 30 per cent when the critical concentration is reached and the 550 m μ peak of tyrosinebisazo-1-H-tetrazole appears. Furthermore, even the availability of the first histidine residue in DIP-chymotrypsin is much lower than in α -chymotrypsin or its substrate complex as the concentration of the diazo reagent must be doubled with respect to the free enzyme in order to complete the reaction. This difference in reactivity between α -chymotrypsin and its DIP derivative seems to be supported by the observation that L-1-tosylamido-2-phenylethylchloromethyl ketone, a specific histidine-modifying reagent, reacts with one histidine residue only in α -chymotrypsin and none in DIP-chymotrypsin (100). It appears, then, that α -chymotrypsin and its BGME substrate complex have two types of histidine residues with different degrees of accessibility, while chymotrypsinogen and DIP-chymotrypsin each have one free and one buried histidine residue. It may be possible, however, that even the more reactive residue of chymotrypsinogen is not completely free, since 4 hr of exposure to pH 13 affect the reactivity of one residue only and 18 hr are needed to render the second residue reactive. The difference in histidine

reactivity between α -chymotrypsin and its zymogen suggests that activation of the zymogen is accompanied by a loosening of the structure around one residue. On the other hand, the formation of DIP-chymotrypsin obviously limits the accessibility of one group. These deductions are supported further by the minimal molar ratio of azo compound to reacted histidines necessary to reach identical degrees of reaction, since this ratio is much higher for DIP-chymotrypsin than for α -chymotrypsin or its substrate complex; the same is true for the chymotrypsinogen- α -chymotrypsin pair. This analysis of the state of histidines in the chymotrypsin system is consistent with other studies. Thus, only one histidine residue in α -chymotrypsin is destroyed by photooxidation (101, 102) and therefore it must be completely exposed, while the second one is buried. Furthermore, only one histidine residue reacts with dinitrofluorobenzene in α -chymotrypsin and DIP-chymotrypsin, while none do in chymotrypsinogen (103).

In cytochrome *c*, three types of histidine residue reactivities have been found (102). In horse heart cytochrome *c* there is one histidine residue of each type, while in yeast cytochrome, the most reactive type is represented by two residues. Reduction of ferricytochrome to the ferrous state decreases the reactivity of histidines only of the second and third types, but shows no effect on that of the first, or most reactive type. The decrease in reactivity upon reduction is much greater for horse heart cytochrome than for the yeast protein. On the basis of these results, it has been concluded that the histidine residues of the first type are completely exposed, while those of the second and third types are of limited accessibility. These results are in complete agreement with titration data on horse heart cytochrome *c* (104, 105), which indicate the presence of only one normal imidazole group, while the other two are assumed to be coordinated to the heme iron atom. In the light of this, it is obvious that the histidine residues of the first type (normal) should not be affected by the oxidation state of the heme iron atom. In its turn, the Soret band of cytochrome *c* should not be affected by the diazotization of these histidine residues. On the other hand, it is not surprising that the reactivity of the histidine residues of the second and third types, which are coordinated to the heme iron atom, should be affected by the oxidation state of that atom, and that diazotization of these residues should affect the Soret band. The histidine residues of the first type in yeast cytochrome *c* show the same degree of reactivity toward the azo reagent as free histidine. This is an additional indication of how unrestricted these residues must be. Comparison of the reactivity of these histidine residues to those of horse heart cytochrome *c* type by type, reveals that all the residues of the horse heart cytochrome are less reactive (less accessible) than their counterparts in yeast cytochrome, indicating a possibly tighter folding of the peptide chain in the horse heart protein. In horse heart cytochrome *c*, histidine 33 was identified as the normal one, while histidines 18 and 26 are linked to the heme iron atom; in yeast cyto-

chrome *c*, histidines 23 and 31 must be the bound ones while residues 38 and 44 are free. Photo-oxidation of horse heart cytochrome *c* (106) results in total inactivation upon destruction of 1.3 moles of histidine, while the kinetics of the reaction indicate that, of three residues, one is destroyed rapidly and a second one slowly.

The histidine residues of ribonuclease react in four steps (93), this difference being eliminated by denaturation. Photo-oxidation studies (106) have shown that at most three histidine residues were destroyed, two being attacked rapidly while the third one was destroyed only after very extensive reaction. Photo-oxidation of ribonuclease S (107) revealed that histidines 105 and 119 are destroyed at approximately equal rates, while histidine 12 is oxidized more slowly, this rate being also slower than that of the oxidation of the same residue in the S-peptide. Considering that the photooxidation studies do not rule out the possibility of a slight difference in reactivity between the two most reactive groups, these results corroborate those obtained with DHT (93), since the discrepancy between the two methods is probably attributable to only a slight gradation in reactivity between the two most reactive groups. Thus, in ribonuclease, the four types of histidine reactivities found between pH 8.8 and 9.8 are reduced to three types between pH 10 and 10.5, and the plateau at one residue disappears from the reaction curve. If, as suggested by the author (93), this decrease in reactivity of the first residue is fortuitous and is only the result of the lower reactivity of the reagent, the observed reactivities of the third and fourth residues have increased. Thus, in the case of ferri and ferro horse heart and yeast cytochromes *c*, a change of pH from 8.8 to 10.6 does not affect the number of reactive histidine residues, but, in complete agreement with the suggested decrease in the reactivity of DHT at the higher pH, all the residues had become apparently less reactive. It would appear, then, that, when the pH is raised to 10.5, the last two histidines of ribonuclease do become more reactive and, thus, probably more exposed as the protein becomes denatured. Similar experiments were carried out on ribonuclease in the presence of nucleotides, namely: cytidine-3'-phosphate (CMP); uridine-3'-phosphate (UMP); adenosine-3'-phosphate (AMP); and guanosine-3'-phosphate (GMP) (93). AMP and GMP had no effect on the histidine reactivities; CMP enhances the reactivity of the residue which is third to react in the absence of the nucleotide, while UMP enhances the reactivities of the second and fourth. Since CMP and UMP are products of the reaction catalyzed by ribonuclease, it has been postulated (93) that the affected histidines take part in the enzyme activity.

In human serum β_1 -lipoprotein (94), about 10 per cent of the histidine residues react with the diazo reagent in both the native and delipidated states. Carboxymethylation with bromoacetate indicates that ca. 50 per cent of the histidines are reactive in the native and somewhat more than 50 per cent in the delipidated state. This case, and that of ribonuclease, indicate

the absence of correlation between DHT reactivity and carboxymethylation.

In soybean trypsin inhibitor, both histidine residues were found to react with DHT (108). Furthermore, on the basis of the kinetics of the imidazole catalyzed hydrolysis of *p*-nitrophenyl acetate, they are considered to be fully exposed to solvent. In carboxypeptidase A (93a), only one of eight histidine residues reacts with DHT, while none are modified in the presence of the competitive inhibitor, β -phenylpropionate.

Tryptophan.—The most widely used method of determining the tryptophan contents of proteins and of cleaving tryptophan peptide bonds is oxidation with N-bromosuccinamide (NBS). Since the subject has been extensively reviewed recently (109), only the latest developments in this field will be considered. As is well known, oxidation by NBS attacks both tryptophan and tyrosine peptide bonds and is, therefore, not a specific reagent for tryptophan. When NBS oxidation is carried out in 8 *M* urea, however, the actual oxidizing agent is N-bromourea (110); its reaction with tyrosine or tyrosine analogues proceeds at a rate >200 times slower than that of NBS in aqueous solution, while the rate of oxidation of tryptophan derivatives is little affected. Thus, the oxidation of phloretic acid gave a 250-fold difference in the rate of tyrosine oxidation, while in lysozyme and insulin, the oxidation rate of tyrosine residues fell drastically. It has been concluded that, in the absence of SH groups, oxidation with NBS in 8 *M* urea is specific for tryptophan residues, if minimal amounts of reagent are used and the tryptophan residues are not buried deeply. A successful method for the cleavage of histidine peptide bonds with NBS has been reported (111). Although it is not possible to cleave the histidine peptide bonds without simultaneously cleaving reactive tryptophan and tyrosine bonds, it is possible to differentiate between the cleavage of these three types. Thus, a His-Pro bond in sperm whale myoglobin was confirmed by this method; furthermore, the authors point out that cleavage of tryptophan and tyrosine peptide bonds does not seem possible without at least partial oxidation of the histidines.

The NBS titration of tryptophans in proteins has given some indications of tryptophan burial in TMV protein (109) and in hemoglobins (112), since some of the tryptophans could be oxidized only after the complete disorganization of the molecular structure in 8 *M* urea. Recently, however, Green & Witkop (113) have reported that the reactivity of tryptophans with NBS is a function of pH, the reactivity decreasing as the pH is raised from four to seven. As most of the proteins which show this effect have isoelectric points above pH 7, it was suggested that a decrease of the pH causes unfolding of the tertiary structure as the net positive charge is increased. None of the examined proteins, however, are known to undergo large conformational changes in this pH range; thus, such effects, if any, must be very small and local. In any case, if this interpretation is correct,

the tryptophan reactivity at pH 4 cannot reflect the accessibility of residues in the fully native state.

In 1964, Hachimori et al. (114) have reported that the states of tryptophan residues in proteins can be investigated by the extent of their oxidation with H_2O_2 in 10 per cent dioxane. Ozonization in the presence of resorcinol has been proposed as a selective oxidizing reagent for tryptophan residues in proteins (115). *N'*-formylkynurenine, the oxidation product, is stable under conditions of acid hydrolysis, rendering this method useful for preserving tryptophan residues during such hydrolysis. Besides tryptophan, only methionine is reported to be oxidized under the conditions used. 2-Hydroxy-5-nitrobenzyl bromide (HNBB) has been introduced by Koshland, Karkhanis & Latham as a tryptophan modifying reagent (116). It is specific for tryptophans in proteins which do not contain free SH groups (117) and has been used successfully for the quantitative determinations of tryptophan residues in lysozyme, α -chymotrypsin, subtilisin, and ribonuclease (118). The results of recent studies on the state of tryptophan residues in proteins, involving NBS, H_2O_2 in 10 per cent dioxane, ozonization, and HNBB are summarized in Table IV.

In B- α -amylase, one out of 14 tryptophan residues was found to react with NBS with no change in conformation (119). In horse heart cytochrome *c*, the single tryptophan residue is oxidized with NBS, while both tryptophan residues react in tuna cytochrome *c* (120). In papain, one tryptophan residue was both NBS reactive and photo-oxidizable (121). In metmyoglobin (122), both tryptophan residues are oxidized by NBS; they are found to be partially, although possibly not equally exposed in solvent perturbation studies, suggesting their presence within crevices. This is in full accord with the known structure of the protein (122a), which reveals that the NH groups of the indole rings are on the surface of the molecule, the remainder of the residue being buried in the interior.

In avidin, which contains 14 tryptophan residues, some of which must be buried as shown by difference spectra of the native protein vs. protein denatured in 8 *M* urea (123), 12 tryptophan residues react with NBS at pH 4 (124). Furthermore, of these 12, ca. 3.5 residues show a higher reactivity (125). At pH 5.5, nine residues react, while at pH 7, only five residues react at a very slow rate; in 8 *M* urea there are 13 reactive residues (113). The formation of the avidin-biotin complex in a 1:3 molar ratio is accompanied by the burial of additional tryptophans as shown by the red shift of the characteristic tryptophan maxima in the spectrum of the avidin-biotin complex (125); this is confirmed by the blue shift of the fluorescence emission spectrum (126), as well as by the fact that ethylene glycol displays a smaller effect on the spectrum of the avidin-biotin complex than on that of biotin alone (125). The polarity of the tryptophan environment falls in the order: denatured > native > complex (126). Burial of tryptophan residues in the complex is accompanied by a decrease in NBS oxidation; no trypto-

TABLE IV
THE STATE OF TRYPTOPHAN RESIDUES IN PROTEINS

THE STATE OF TRYPTOPHAN RESIDUES IN PROTEINS

Protein	Tryptophan, total	Reactive with							Ozone	Solvent perturbation	
		NBS			H ₂ O ₂ , 10% dioxane		HNBB				
		pH 4	pH 5.5	Urea	Native	Denat.	Native	Urea			
β - α -amylase	12	1									
Cytochrome c, tuna	2	2									
Cytochrome c, horse heart	1	1									
Papain	5	1									
Metmyoglobin	2	2									
Avidin	14	12(2types)	9	13							2 partly exp.
Avidin-biotin complex	14	0									
Pepsin	6	3+1	2.4	6					4+1		
Pepsinogen	6	1.8	1.2						4		
Gramicidin	4	4		4							3 exp; 4 buried
Chymotrypsinogen	7	7	4		5+1+1						2 exp; 1 partly exp;
α -Chymotrypsin	7	7	3	7	5+1	7	1-3	8.1	4+1		4 buried
α -Ct-BGEE	7				5+1						same at Ct
DIP-chymotrypsin	7			6.6	4+2						all partly exp.
Lysozyme	6	6	3.5	6	5+1		1.5	5.6			
Trypsinogen	4	4	2		1+1+1	4					
Trypsin	4	4	2	4	2+1	4		4.2	3+1		2 partly or fully exp.
Soybean trypsin inhibitor	3	1	1	3	1+1			3			3 exp.

phans are oxidized in the avidin-biotin complex after short exposure to NBS. It is only after a 30 minute exposure and at a higher NBS/protein molar ratio (relative to that for avidin) that 30 per cent of the residues are induced to react.

In pepsin (127), four tryptophans can be oxidized with NBS at pH 4, while the other two become accessible only in 8 *M* urea. Furthermore, it appears that the four reactive residues do not have equal rates of oxidation, the first three oxidizing much faster than the fourth. Green & Witkop (113) find that 3.6² tryptophans oxidize at pH 4 while only 2.4² react at pH 5.5 and pH 7 with a decreased rate. In pepsinogen, 1.8² tryptophan residues react at pH 4, while at pH 5.5 and 7, the reactivity falls down to 1.2² (113). Ozonization of pepsinogen results in the oxidation of 4.7 tryptophans (130), with four residues oxidizing faster than the fifth one. In gramicidin, four tryptophan residues are found to react both with NBS (113) and ozone (131).

In chymotrypsinogen, the seven tryptophan residues are oxidized with NBS at pH 7 (132). On the basis of oxidation with H₂O₂ in 10 per cent dioxane, the tryptophan residues can be divided into three types consisting of five, one and one residues each (114). In α -chymotrypsin, seven tryptophans are oxidized in NBS at pH 4 and in 8 *M* urea (113), with two residues being more reactive than the rest (133). At pH 5.5 the reactivity decreased to four groups for the zymogen and three for the enzyme (113, 134). The reactivity of α -chymotrypsin tryptophan residues toward H₂O₂ in 10 per cent dioxane was found to be of two types, consisting of five and one residues each (114). After denaturation at pH 12.5, however, all seven residues underwent the reaction. These results suggest that the reactivity of the tryptophans is related to the secondary and tertiary structures; furthermore, from the minimal H₂O₂ concentration required for the oxidation of the most reactive five residues (relative to N-acetyltryptophan), it has been concluded that even these residues are not fully exposed. Reactivity with HNBB resulted in one to three (126) and three (135) reactive residues in the native state, while in 8 *M* urea 8.1 tryptophans were reported to be affected (118). Ozonization of α -chymotrypsin resulted in the loss of 4.5 tryptophans (130), with four residues apparently reacting at a faster rate than the fifth one. Photo-oxidation leads to the destruction of three residues (101), while enzymatic oxidation with horse radish peroxidase results in the loss of only one (136). The difference in reactivity of one tryptophan residue between α -chymotrypsin and its zymogen, as shown both by NBS (pH 5.5) and H₂O₂ oxidation, would suggest that the activation of the zymogen is accompanied by a tightening of the structure about one tryptophan. This is confirmed by the results of solvent perturbation studies which

² The results of Green & Witkop were reported on the basis of five tryptophan residues per molecule of pepsin and pepsinogen; these have been recalculated on the basis of six (128, 129).

indicate that, of the three fully exposed tryptophans in chymotrypsinogen, one residue becomes partially buried after activation (132, 137); the activation of the zymogen to enzyme did not affect the degree of exposure to long range perturbants, only access to short range perturbants was decreased (137), suggesting that the environment of some exposed residues had been altered during activation. In view of these reports, the higher reactivity of two tryptophan residues toward NBS in α -chymotrypsin (133) can be interpreted in terms of the oxidation of fully exposed groups. The tryptophan residues of the α -chymotrypsin-BGME complex have the same reactivity with H_2O_2 in 10 per cent dioxane as the uncomplexed protein (139). In DIP-chymotrypsin, the reactivity toward H_2O_2 changes from two types, of five and one residues each in the native protein, to two, of four and two residues each (91); furthermore, from the minimal amount of H_2O_2 required, it appears that the reactivity of the least reactive residue is depressed even more relative to α -chymotrypsin. This leads to the conclusion that phosphorylation of the native protein results in a tightening of the structure about two tryptophans. The difference between the reactivity to H_2O_2 of the tryptophans of α -chymotrypsin and DIP-chymotrypsin must be related to those residues which are buried both in α -chymotrypsin and in the derivative, since solvent perturbation shows that, on reaction, the extent of exposure does not change substantially (137). This is supported by the conclusion of Oppenheimer, Mercouroff & Hess (1933) that the number of exposed tryptophans is the same in α -chymotrypsin as its DIP-derivative and that their observed difference spectra might be assigned best to those residues which are buried in both molecules. The difference in reactivity of the α -chymotrypsin and DIP-chymotrypsin tryptophans persists even in 8 *M* urea, the α -chymotrypsin tryptophans being more susceptible to NBS oxidation than those of DIP-chymotrypsin (139).

In lysozyme, six tryptophan residues are oxidized with NBS at pH 4 (113, 132, 140); at pH 5.4, only 3.5 residues react at a lower rate (113). Comparison of the oxidation in buffer and in 5 *M* urea shows that the first residue is oxidized at about the same rate in both media, while the next three are oxidized more rapidly in urea. This may either reflect the secondary and tertiary structures of the protein or simply the consumption of NBS by $-\text{SH}$, $-\text{S}-\text{S}$ and free $-\text{NH}_2$ groups (109). Oxidation with H_2O_2 in 10 per cent dioxane (114) leads to the conclusion that there are two classes of tryptophan residues, consisting of five and one residue, respectively. From the minimal H_2O_2 concentration requirements, it is concluded that some of the five more reactive residues are not completely accessible to the reagent; the last one is the least accessible of the six. In the case of HNBB, 1.5, tryptophan residues react in the native state (135) and 5.6 residues react in 8 *M* urea (118). Photo-oxidation destroys three residues (97). These results at first appear to be somewhat confusing. They are clarified, however, by the solvent perturbation studies which have shown that all six

residues are exposed to solvent to various degrees and are not located within crevices (132, 137), a conclusion quite consistent with the known structure of this protein. It is reasonable that such a partial exposure will prevent clear-cut gradations in the reactivity of the individual residues and result in overlapping classes with different reagents. One residue appears to be particularly reactive, as suggested by modification with HNBB (135), as well as by iodination studies (141). Oxidation with NBS at pH 5.5 (a not very selective reagent) may reflect the same level of accessibility as photooxidation, which is probably a rather selective method; oxidation with H_2O_2 also suggests a partial exposure of the residues. A red shift (142), believed to be related to a change in environment of one tryptophan residue, is observed in the spectrum of lysozyme when that protein forms a complex with glycol chitin. The residue in question must be the most easily oxidizable one, since the difference spectrum disappears if one tryptophan is oxidized prior to complex formation (142). Solvent perturbation studies (143) have shown that complex formation decreased the exposure of tryptophans to sucrose from six groups to 3.5, while exposure to polyethylene glycol (PEG) 400 changed from 3.8 to 2.7 residues. Thus, the complex formation affects 42 per cent (2.5 groups) of the tryptophans when judged by their exposure to sucrose, and only 18 per cent (1.1 group) according to their exposure to PEG. Just as in free lysozyme, the exposure of the tryptophans in the complex varies, depending on whether sucrose or PEG 400 is used as perturbant, although the difference is not as large (2.2 groups in lysozyme, 0.8 in the complex). The authors (143) have interpreted these results in terms of completely buried and completely exposed residues, neglecting the possibility of partial exposure (132, 137, 144). Considering the fact that a decrease in the extent of exposure on going from sucrose to PEG indicates partial exposure (144), this must be the case in the complex, as well as in the native lysozyme. Therefore, the formation of the complex must be accompanied by a change in the environment of more than one tryptophan residue, since between 1.1 and 2.5 are affected, if interpreted on the all buried or all exposed basis; the more plausible interpretation of a change in the degree of burial leads to the conclusion that even more tryptophan groups are involved.

In trypsinogen, four tryptophan residues react with NBS at pH 4, and two at pH 5.5 (113, 145). Three reactive residues are found on oxidation with H_2O_2 in 10 per cent dioxane (114). After denaturation, four groups react. From the minimal H_2O_2 concentration requirements, it can be assumed that these residues differ in their degree of exposure. In trypsin, four residues react with NBS at pH 4, two at pH 5.5, one at pH 7 (at a lower rate) and four in 8 *M* urea (113, 145). There are three H_2O_2 reactive residues in the native state and four in the denatured state (114). The three residues do not seem to differ from those in trypsinogen, except possibly for a very slight difference in the reactivity of the second one. Thus,

on the basis of oxidation with H_2O_2 , it would appear that activation of the zymogen does not affect any tryptophan residues, while the reactivity of histidine and tyrosine residues becomes greatly enhanced (90, 146, 147). This result (114) is in disagreement with Viswanatha, Lawson & Witkop (145) who reported that the tryptophan residues are more reactive with NBS in trypsinogen than in trypsin, and concluded that activation of the zymogen is accompanied by a tightening of structure in the environment of some tryptophan residue(s). Furthermore, difference spectra observed upon activation of the zymogen (148) suggest that both tyrosines and tryptophans are affected by the reaction. All the residues react with HNBB in 8

TABLE V
STATE OF TRYPTOPHAN AND TYROSINE RESIDUES IN
TRYPSIN—INHIBITOR COMPLEXES

Inhibitor in complex	Tryptophans				Tyrosines
	NBS Reactive			Shielded in complexed trypsin	Shielded in complex
	In trypsin	In inhibitor	In complex		
Pancreatic	3	0	2	1	3
Lima bean	3	0	1	2	3
Ovomucoid	3	0	2	1	2
Soybean	3	1	2	1	3

M urea (118). Ozonization leads to the loss of 3.5 residues (130), three residues oxidizing at an identical rate, while the last one is less reactive.

In soybean trypsin inhibitor, one tryptophan reacts with NBS at pH 4 to 7 (108, 109) while three react in 10 *M* urea; two residues are oxidized with H_2O_2 in 10 per cent dioxane with some indication that they are not identical (108); three residues react with HNBB in 10 *M* urea (108); solvent perturbation studies show that at least two tryptophans are partially or completely exposed (108). An interesting recent report (149–151) describes changes in the reactivity of both tryptophan and tyrosine residues as well as in physical properties when trypsin complexes with several inhibitors. While NBS oxidation of trypsin at pH 5.2 results in the destruction of three tryptophan residues, formation of a 1:1 complex with pancreatic (PTI), lima bean (LBI), or soybean (SBI) inhibitors, as well as ovomucoid, leads to a decrease in the reactivity of the tryptophans. The results, summarized in Table V, point to the shielding from reaction of both tryptophan and tyrosine residues. In the STI-trypsin complex only one of the shielded tryptophans is assumed to be present in trypsin; the second one may be in the inhibitor, since STI contains three tryptophans, one of

which is susceptible to NBS oxidation. The other three inhibitors do not contain tryptophan, so that the tryptophans affected must be located in the enzyme. Since all the inhibitors contain tyrosine, it is impossible to decide where the shielded residues are. In the PTI complex, they appear to be in the enzyme. All trypsin-inhibitor complexes develop difference spectra (red shift), which are similar qualitatively, although different quantitatively. The wavelengths are characteristic both of tyrosyl and tryptophyl groups, confirming that complex formation alters the environment of both.

Tyrosine.—The relation between anomalous tyrosine ionization and the secondary and tertiary structure of proteins was first observed by Crammer & Neuberger (152) in ovalbumin by means of spectrophotometric titration. Since then, the pK , ΔH and ΔS of tyrosine dissociation have been widely used to probe the effect of protein structure on individual tyrosine residues built into protein molecules. By the criterion of titration, all irreversibly titrating tyrosines are buried, while, among the reversibly titrating ones, only those which are similar to small tyrosine peptides in their thermodynamic properties ($pK_{int} = 9.6$, $\Delta H^\circ_{int} = 6$ Kcal/mole, $\Delta S^\circ_{int} = -24$ cal/mole degree) are regarded as unquestionably normal (96). All the other reversibly titrating groups, although obviously accessible to hydrogen ions, are not necessarily normal (completely exposed to the solvent). These tyrosine residues are classified as being of intermediate exposure, i.e., neither completely exposed nor completely buried.

No specific tyrosine modifying reagents seem to have been available until quite recently when reports that N-acetylimidazole (153) and cyanuric fluoride (154) react with tyrosine residues opened the possibility of probing these groups chemically. When reacted with tyrosine residues, these reagents give O-derivatives in which the π - π^* transition responsible for the 2750 Å absorption band is strongly affected. The changes in ultraviolet absorption which are caused by these transformations can be observed spectrophotometrically and used for the quantitative estimation of the number of affected tyrosine residues. Thus, in the case of O-acetylation of N-acetyl-tyrosine, the maximum shifts from 275 to 263 m μ (153); a similar effect has been observed for tyrosine-O-sulfate present in bovine fibrinogen (155).

N-acetylimidazole is considered to be a diagnostic reagent for free tyrosine residues (153) while cyanuric fluoride is a reagent applicable to any reversibly titrating tyrosine (154). The states of tyrosine residues in myoglobin (153), ovomucoid (153), rabbit muscle aldolase (153), BSA (153), carboxypeptidase A (153), iron conalbumin (153), liver alcohol dehydrogenase (153), trypsin (153), α -chymotrypsin (153), insulin (153), ribonuclease (153), ovalbumin (153), hemoglobin (153), conalbumin (156), cytochrome *c* (157), pepsinogen (158), pepsin (158), lysozyme (159), α -lactalbumin (159), β -lactoglobulin (160), and soybean trypsin inhibitor (161) have been studied using N-acetylimidazole. Similar studies using cyanuric fluoride (CyF) have been carried out on α -chymotrypsin (91),

TABLE VI
THE STATE OF TYROSINE RESIDUES IN PROTEINS

Protein	Titration			N-Ac-Imid.		CyF	TNM	Solvent perturbation
	pK ^a	Number		Native	Denat.			
		Rev.	Irr.					
Myoglobin	10.3, 11.5, 12.8	1+1	1	0.8				
Ovomucoid		5		5.3				most exp. to small, but not to bulky perturbants
Rabbit muscle aldolase	10.4	11-13	31-33	10.0				14 gps. exp.
Bovine serum albumin	10.3	18		4.0				50% acc. to small, 30% acc. to bulky
Carboxypeptidase A	9.5	7	8	5.9	19		7.1	
	9.4	8	8-9					
Conalbumin, Fe		5	13	5.0				
Conalbumin	9.4	11	7	10-11				
Liver alc. dehydrogenase			10 (total)	6.1				
Trypsin	10.0, 10.8	6	4	6.7				
α-Chymotrypsin	10.2, 11.3, 12.5,—	1+1	1+1	2.0	4	2		(40 to 75% exp. in chymo- trypsinogen)
DIP-chymotrypsin	10.2, 11.3, 12.5,—	1+1	1+1			1		

^a The pK values reported are either intrinsic or apparent, depending on the value given in the original publication.

TABLE VI (Continued)

Protein	Titration			N-Ac-Imid.		CyF	TNM	Solvent perturbation
	pK ^a	Number Rev.	Irr.	Native	Denat.			
α -Ct-BGME						1		
Insulin	9.6	4		4.1	4	2+1+1 or (2+1)	2	Zn: 2 exp.; Zn-free: 3 exp.
Ribonuclease	10.4, 11.4	3	1					
	9.9	3	3	3.0	6	1+1+1	3.1	50% exp. (4 partly exp., 2 buried)
Ovalbumin	11.4-11.8	2 (?)	7-8	1.5		2	5.6	
Hemoglobin	10.6, >12	8	4	7.6				
Cytochrome <i>c</i> , Fe III	10.0, 11.0 12.35, 13.1	1+1	1+1	2	4			all buried
Cytochrome <i>c</i> , Fe II				>0.5				
Pepsinogen	11.0	15	2	8-10			10.4	
Pepsin	10.0, 11.5 >11.7	11	5+(1-2)	9.2				
Lysozyme	10.8	3		2		2		50%, or 1 or 2 groups exp.
	10.5, 12.8	2	1					
α -Lactalbumin	10.4	4	1	4		3+1		
β -Lactoglobulin	9.9, 10.9, 12.2	3	1	4	4	2+1		
Stem Bromelain	9.8, 11.8, <12, <13	9	7+1+2			(7-8)+ (5-7)		
Oxytocin	10.0	1				1		
Glucagon	10.0	2				1+1		

insulin (154, 160), lysozyme (154), DIP-chymotrypsin (91), the α -chymotrypsin—BGME complex (91), ribonuclease (160), α -lactalbumin (160), β -lactoglobulin (160), ovalbumin (161), stem bromelain (162), oxytocin (163), glucogen (163) and soybean trypsin inhibitor (161). Table VI contains the results of these studies and compares them with pertinent literature data.

The reactivity toward N-acetylimidazole agrees with the titration data on the number of normal tyrosine residues in 12 proteins out of the 21 reported cases. Furthermore, in the case of pepsin, the discrepancy between reactivity toward N-acetylimidazole and titration (164) might be fortuitous, since the acetylation reaction had to be carried out at pH 5.6 to 5.8 to preclude the denaturation of the protein. At this pH, however, N-acetylimidazole is not stable, with the probable result that complete acetylation cannot be accomplished. If a discrepancy exists between the number of normal tyrosine residues determined by the two methods, the protein may be placed in one of two classes. The first class, in which N-acetylimidazole indicates fewer normal residues than titration, contains rabbit muscle aldolase (165), carboxypeptidase A (166, 167), and pepsinogen (168). Trypsin (169), β -lactoglobulin (160), and BSA (96) belong to the second class in which the chemical reagent indicates more normal groups than titration. BSA is particularly interesting: although all the tyrosine residues titrate reversibly with an only slightly elevated pK, the ΔH and ΔS values are unusually high (11.5 Kcal/mole; -9 cal/mole degree), indicating that, in spite of their accessibility to hydrogen ions, these tyrosines cannot be completely free. It is tempting to explain the discrepancy between the two methods in terms of a lower sensitivity of the titration technique to the environment of tyrosines in proteins; this difference might be related to the nominal spatial requirements of hydrogen ions and of the N-acetylimidazole reagent. In terms of this interpretation, any case in which the number of groups which react with the reagent is less than the number of normally titrating ones (rabbit muscle aldolase, carboxypeptidase A, pepsinogen) or of identically titrating ones (BSA) would serve to demonstrate the greater discriminatory power of N-acetylimidazole. At present, there is very little evidence to substantiate such a claim. Solvent perturbation spectroscopy (144), which is probably the most sensitive method for probing into the environment of chromophoric residues in proteins, offers only very few examples for comparison.

As can be seen from Table VI, lysozyme and ribonuclease, two proteins for which N-acetylimidazole reactivity and titration (170) give identical results, show solvent perturbation data in agreement with these findings only in the lysozyme case (171). In ribonuclease, solvent perturbation indicates that the three normal groups are, in fact, not identical (172). In insulin, according to the solvent perturbation results, the tyrosine residues can be grouped into three types (2+1+1), the second type appearing only on

binding of zinc (173). (In Zn-free insulin, three groups are exposed and only one buried.) These differences are discerned neither by N-acetylimidazole nor by titration (96), both of which indicated four normal residues, although alternate titration results of three normal and one buried groups (170) have been reported. In cytochrome *c*, N-acetylimidazole reactivity and one set of titration data (178) agree on two normal tyrosine residues, while solvent perturbation indicates that all four residues are buried (174), as do the results of photo-oxidation experiments (175) which gives a loss of no more than 10 per cent of the residues. Aldolase gives no correlation of any sort, while BSA provides an example in which the sensitivity of N-acetylimidazole appears to be equal to that of solvent perturbation (176).

Whether N-acetylimidazole is a diagnostic reagent for normal tyrosine residues has not been established as yet. It should be pointed out that it reacts with two tyrosine residues in α -chymotrypsin, the second one of which cannot be considered normal because of its high pK of 11.3 (96). The same might be true also for cytochrome *c*, for which one (177), two (178) and zero (174) tyrosine residues have been reported to titrate normally. Furthermore, in trypsin (153), β -lactoglobulin (160), and possibly insulin (153), it can react with buried tyrosine residues.

The number of studies with cyanuric fluoride reported up to now is rather small. From those compiled in Table VI, it would appear that CyF is capable of interacting with probably any type of reversibly titrating residue. Furthermore, in several cases, where comparative data are available, it seems capable of discerning gradations in the reactivities of tyrosine residues more subtle than those discernible by titration or by the N-acetylimidazole reagent. Thus, in α -chymotrypsin (91), ovalbumin (161), lysozyme (154), and oxytocin (163), the CyF data are in complete agreement with the titration and N-acetylimidazole results. Interestingly enough, in DIP-chymotrypsin (91), which titrates the same as α -chymotrypsin, and in the α -chymotrypsin—BGME complex, only one tyrosine reacts with CyF. In stem bromelain (162), for which N-acetylimidazole data are lacking, CyF reveals two types of tyrosines: out of a total of 19 groups, 13 to 14 react: of these, 7 to 8 show a much higher reactivity than the rest; titration data indicate nine normal residues with pK = 9.84 and seven supposedly buried ones with pK = 11.8 (the other three have higher pK's and are still less accessible). The agreement between the two methods is quite good although not completely quantitative. In the case of soybean trypsin inhibitor, all four tyrosines titrate normally with a pK of 9.5 (179), while solvent perturbation gives 70 per cent of the groups exposed (108) and the iodination reaction indicates the presence of two reactive groups at pH 8 to 9 and four in 9 *M* urea (108). Reaction with N-acetylimidazole results in three normal residues (161), and CyF reacts with one group at pH 9.5 and two groups at pH 11.5 (161). Comparison of these various techniques would suggest that all four groups lie in somewhat different environments.

The high sensitivity of CyF to tyrosine residue environment is seen in ribonuclease, α -lactalbumin, insulin and β -lactoglobulin. In ribonuclease (160), CyF reveals three types of normal tyrosines as contrasted with only one type indicated by titration and N-acetylimidazole reactivity. In α -lactalbumin (160) which, according to titration (160) and N-acetylimidazole reactivity (159), has four normal residues, CyF shows that three of them are more reactive than the fourth. In insulin (154, 160), where the number of normal tyrosines is four according to N-acetylimidazole reactivity, while by titration, it is either four (96) or three (170), CyF indicates three reactive groups (160), of which two are more reactive than the third. In β -lactoglobulin, for which titration indicates three normal and one buried tyrosine residues (96, 160) and N-acetylimidazole gives four identical ones (160), CyF reacts only with three, of which again two are more reactive than the third.

The discriminatory power of CyF receives further support from solvent perturbation studies. Thus, for lysozyme, the solvent perturbation results confirm the CyF reactivity of two groups. In Zn insulin, solvent perturbation shows that two groups are exposed and two buried, but probably in different ways, since a third group becomes exposed on removal of zinc. The gradation in CyF reactivity of 2+1+1 reactive tyrosines seems to fall into the same pattern as that of exposure to perturbant. Solvent perturbation studies on ribonuclease indicate that, in aqueous solutions, somewhat more than one half of the tyrosine residues are unavailable to various perturbants, while in non-aqueous solutions only two tyrosine residues are buried, the other four being accessible to solvents to varying degrees (171, 172). The results of these studies imply that the normal tyrosine residues of ribonuclease are not, in fact, identical with respect to their environment, an implication in complete agreement with the CyF reactivity data, as well as with observed differences in the rates of iodination (180) and the fact that only one residue is oxidized with tyrosinase (181).

Quite recently tetranitromethane (TNM) (182, 183) has been reported to be a tyrosine modifying reagent. The results, summarized in Table VI, show that, in the case of insulin and ribonuclease, the number of groups which react in the native protein is equal to that found normal by the CyF reaction. In carboxypeptidase A, the TNM value of seven reactive groups agrees with the number of reversibly titrating tyrosines and is one greater than shown by reactivity with N-acetylimidazole, while the reactivity with pepsinogen is equal to that of N-acetylimidazole but considerably below the number of reversibly titrating groups. Ovalbumin is one protein in which reactivity with TNM leads to a considerably larger number of accessible groups than shown by any of the other techniques.

The similarities and differences between the number of tyrosine residues accessible to the various reagents points to the complicated interplay of factors that define the "exposure" or "burial" of any group; these may involve dimensions and geometry of crevices, charge configuration, pres-

ence of specific vicinal groups, steric requirements of the transition state and minor local conformational changes that may occur during reaction. It is hard to expect that all the methods would give identical results; on the contrary, one might expect the opposite to be true. By a comparison of the various methods, however, it should become possible to arrive at conclusions on the nature of the environment of a given type of group. In order to accomplish this, however, it will be necessary to accumulate far more comparative data than are available at present.

Solvent perturbation spectroscopy.—One of the most sensitive and highly discriminating techniques available for examining the location of chromophoric residues with respect to the surface of a protein molecule is that of solvent perturbation spectroscopy, developed by Laskowski and co-workers (144). This technique is based on the fact that the ultraviolet absorption spectra of aromatic compounds are considerably influenced by the medium in which these substances are dissolved. The exposure of proteins to solvents of higher refractive index than water results in a small change in the extinction coefficient and a small red shift of the absorption maximum. The magnitude of these changes depends on the location of the chromophores within the framework of the protein molecule. Thus, chromophores which are located close to the surface and capable of coming in contact with the molecules of the solvent will suffer characteristic red shifts. On the other hand, chromophores which are deeply buried in the interior of the molecule and are not capable of interacting with the solvent will not undergo such shifts. Intermediate cases are given by groups which are present within crevices in the structure of the protein. Such groups will appear to be on the surface when the probing solvent is small and can penetrate into the crevice and come in contact with the chromophore; when the probing solvent is not capable of such penetration, the same group will appear to be buried. In the solvent perturbation technique, a difference spectrum is measured between a protein in a single solvent and the same protein dissolved in a medium to which a perturbing agent, such as methanol, glucose or glycerol, had been added. Analysis of the effect of perturbants of different sizes on the spectra of proteins leads to a classification of the chromophoric residues into buried, exposed, and those present in crevices of various dimensions and shapes. In the present review, neither the theoretical foundations nor the experimental technique as such will be discussed, since adequate reviews are available (144, 184–186). Those results which can be compared directly with the chemical studies have been summarized in Tables IV and VI. In this section, a few most recent reports will be discussed.

In an interesting recent study, the state of tyrosine residues in paramyosin has been followed through the helix-coil transition (187). At pH 7.2, $\Gamma/2 = 0.6$, solvent perturbation with 20 per cent polyhydroxy perturbants of increasing size has shown that of the 38 tyrosines present in the molecule,

35 per cent are totally buried and 10 to 15 per cent are located in crevices, since they are accessible only to perturbants with diameters smaller than 5Å. The rest must be completely exposed. Helix-coil transition, induced by increasing concentrations of guanidine hydrochloride, was found to take place in four steps. During the first two steps, 68 per cent of the tyrosine residues become accessible to solvent, while the last two steps, when the final breakdown of helical structure takes place, release an additional 12 residues to contact with solvent. These are believed to be the residues that are inaccessible to all perturbants. Riddiford & Scheraga (188) have reported that in 0.3 M KCl, 15 per cent of the tyrosines were inaccessible to hydrogen ions. At $\Gamma/2 = 0.6$, all the tyrosine residues titrate normally (189, 190). These results must be considered in the light of studies on the state of aggregation of paramyosin (191) which appears to be a mixture of monomers, dimers, and trimers, depending on ionic strength. At pH 7.2, $\Gamma/2 = 0.6$, this protein exists as a monomer; thus, the solvent perturbation results are not a reflection of group burial on aggregation. The titration results (189, 190), however, also are for a nonaggregated protein, so that paramyosin appears to contain a number of tyrosine groups which are accessible to hydrogen ions, but not to somewhat larger molecules.

In ovomucoid, exposure is found to depend on the size of the perturbant (192), indicating that the tyrosine residues are located within crevices. Most of the tyrosine residues are accessible to small perturbants at pH 4-8. Below pH 4, a large fraction of tyrosine residues become buried as the molecule undergoes a transformation.

In β -lactoglobulin, more than half of the two tryptophan residues per chain are inaccessible to sucrose, ethylene glycol, and dimethyl-sulfoxide between pH 2 and 7 (171). Since, below pH 3, the β -lactoglobulins dissociate into two identical subunits (193), this result indicates that neither of the tryptophan residues is located on the inter-subunit surface. These results can be interpreted in terms of either one residue being fully and the second partially buried, or both being partially buried (194). Furthermore, since the solvent perturbation spectrum is not affected (194) by the low temperature tetramerization of β -lactoglobulin A (195), neither tryptophan of either chain can be located in the site of that specific interaction.

In α -lactalbumin (196), two tryptophan residues are exposed and three buried at pH 1.8-6. At 1° and pH 6, these two exposed groups become inaccessible to sucrose and glycerol, but not to heavy water, while at pH 1.8-3, 1° only one group shows such behavior. The dependence on the size of the perturbant indicates that the chromophore must be located in a crevice. Thus, the two exposed tryptophan residues are located in crevices the size of which is controlled by temperature and pH.

These four examples demonstrate the applicability of solvent perturbation to studies on intermolecular interactions or conformational changes, since in them the macromolecular changes are related to the location within the molecule of the chromophoric residues.

In aldolase (197), one tryptophan and 14 tyrosine residues are found to be exposed to 20 per cent ethylene glycol at pH 6. At pH 2, after a conformational change had taken place, 10 tryptophans and 42 tyrosines become accessible, indicating the complete disorganization of the molecule; these results are in reasonable agreement with the titration properties of the tyrosines (198) and their reactivity with N-acetylimidazole (153). In human carbonic anhydrase B, one third of the tryptophan and tyrosine residues are found to be exposed to sucrose at pH 7 (199), which means that no more than two tryptophans and two tyrosines can be completely accessible to solvent. Difference spectral studies at acid pH's suggest that most or all of the tryptophan residues are buried (200, 201), while titration data (202, 203) indicate the presence of two normal and three reversibly titrating tyrosine residues, in full accord with the solvent perturbation results.

Highly detailed solvent perturbation studies have been carried out on the α (undissociated) form of L-glutamate dehydrogenase (204). These show that 22 per cent of the tryptophan, 58 per cent of the tyrosine and 75 per cent of the phenylalanine residues are buried; the number of completely exposed residues is 40 per cent of the tryptophans, 23 per cent of the tyrosines and 9 per cent of the phenylalanines. Furthermore, the authors state that 39 per cent of the tryptophans are located in crevices, 2.24 to 4.30 Å in diameter, while only 9.5 per cent of the tyrosines and 6 per cent of the phenylalanines are in crevices. The distributions of tyrosine and phenylalanine residues in holes ranging in size from 2.24 Å to 9.74 Å (from totally buried to totally exposed) shows a one to one correspondence. The most interesting feature of this report is that most (75 per cent) phenylalanine residues and two thirds of the tyrosine residues are buried, while only 22 per cent of the tryptophans are buried. The tryptophans are found to be nonrandomly distributed in three discrete "regions" of the enzyme, while neither tyrosines nor phenylalanines show such a pattern. It should be borne in mind, however, that all these results have been interpreted (204) in terms of fully exposed or fully buried residues neglecting the possibility of partial exposure. (For example, 20 per cent of totally buried groups could also be 40 per cent of half exposed groups.) Dissociation of the enzyme into three B subunits results in no change in the environment of chromophoric residues (205), indicating that none of the groups are located in the surfaces of subunit contact; the binding of L-glutamate to the enzyme perturbs one tryptophan residue (206) which, the authors conclude, must be located close enough to the binding site to interact with the bound glutamate.

Another study of an enzyme-substrate complex has been carried out (207) on the trans-cinnamoyl- α -chymotrypsin complex. It is found that, once the complex is formed, the ultraviolet absorption of the cinnamoyl group cannot be perturbed even with such a nonbulky solvent as ethylene glycol. Disruption of the structure by peptic digestion at pH 2.0 does result in a solvent perturbation spectrum. It is concluded that the added chromo-

phore becomes buried in the protein molecule as a result of a conformational change which occurs during the formation of the bond between the enzyme and the substrate.

Lysozyme, which has been subjected to a detailed solvent perturbation examination, presents an interesting example of how differences in interpretation of the data can lead to quite divergent conclusions. As stated above, the solvent perturbation and NBS oxidation studies of Williams & Laskowski (132) and of Williams, Herskovits & Laskowski (137) have led to the conclusion that all the tryptophans are partly exposed. Hamaguchi and co-workers (208-215), on the other hand, have concluded that four tryptophan residues are fully exposed and two are fully buried. Since these workers based their interpretation on the concept of fully exposed vs. fully buried groups, neglecting the possibility of partial exposure (144), the difference between the two conclusions reflects a difference in interpretation and not in experimental results. The tyrosine residues in lysozyme have been variously reported to be accessible to solvent to the extent of two thirds of the total (171), one residue (213) and 45 to 40 per cent (132). The last value refers to exposure to PEG and ethylene glycol and was obtained on a heavily modified (NBS oxidized) protein. The known chemical evidence (see above) is consistent with situations in which two tyrosine residues are fully exposed (two thirds of the total), or exposed to the extent of 50 per cent, or in which one residue is fully exposed and one half exposed (45 to 50 per cent). Thus, the various interpretations of the solvent perturbation data are essentially consistent with the chemical evidence, and it is difficult to choose between them without referring to the complete structure of the protein.

Conclusion.—The comparison, presented above, between the results of various chemical and spectroscopic methods of examining specific groups in proteins shows how such a concerted analysis can lead to details of protein structure in solution which might otherwise be missed.

One further remark seems essential: the results obtained from chemical modification studies can have great significance for native proteins only if the conformation is not changed. Unfortunately, so far such controls have been carried out as an exception rather than as a rule. In the case of β -lactoglobulin, circular dichroism experiments have shown (36) that reaction with N-acetylimidazole results in a small conformational change which, however, is still within the reversible range, and during which the mutual ordering of tryptophan residues has not been affected. In the case of ribonuclease, acetylation of three tyrosines with N-acetylimidazole, which is accompanied by the simultaneous blocking of 79 per cent of the lysines, does not induce any apparent changes in the conformation of the protein; this is shown by a lack of change of the ORD spectrum both in the region of rotational bands attributable to aromatic chromophores (260 to 300 m μ) and in the region reflecting conformations of peptide bonds (216). A particularly interesting ORD study was carried out on cytochrome *c* and its chemical

derivatives (120). In this study, acetylation of two tyrosines with N-acetylimidazole does not alter the Cotton effects having peaks at 278 and 287 m μ , showing that these peaks do not arise from the "free" tyrosyl groups. This observation can serve, as well, as evidence that acetylation does not disrupt drastically the structure of the protein. On the other hand, NBS oxidation of the single tryptophan obliterates the peak at 278 m μ , showing that this oxidation-reduction-dependent band is related to this residue; the lack of any concomitant change in the 287 m μ band suggests again that native structure is preserved elsewhere in the molecule. Similarly, the blocking with N-acetylimidazole of three tyrosines in soybean trypsin inhibitor does not affect the CD spectrum of this protein either in the aromatic or the conformational regions (161). Evidence of this nature that the chemical modification in question does not induce any serious alterations in the structure of the protein is quite essential, if the results are to be associated with the numbers of exposed groups and their relative degrees of exposure. This remark seems even more pertinent when it is realized that essentially none of the "specific" reagents react exclusively with one type of residue. It is sufficient to cite the case of NBS which oxidizes tyrosines and histidines as well as tryptophans, or that of N-acetylimidazole which also reacts with a number of basic residues in addition to tyrosine, example, in ribonuclease 79 per cent of the lysines become blocked concomitantly with the reaction of 50 per cent of the tyrosines (216).

In view of these serious considerations, one might expect in the future to see more comparative studies using a number of reagents with proper controls to prove the absence of conformational changes during the chemical reaction. Simultaneously, it is to be hoped that the necessary spectroscopic techniques will be constantly improved to give more resolution and discrimination. Finally, one can look forward to the development of new physical approaches to this problem: one particularly promising new development is the recently proposed technique for the measurement of the volume of exclusion (217) as a probe of the detailed topology of a protein molecule.

LITERATURE CITED

1. Schellman, J. A., and Schellman, C., in *The Proteins*, 2nd ed., 1-137 (Neurath, H., Ed., Academic Press, New York, 1964)
2. Harrington, W. F., Josephs, R., and Segal, D. M., *Ann. Rev. Biochem.*, **35**, 599 (1966)
3. Urnes, P., and Doty, P., *Advan. Protein Chem.*, **16**, 401 (1961)
4. Fasman, G. D., in *Methods in Enzymology*, **VI**, 928 (Colowick, S. P., and Kaplan, N. O., Eds., Academic Press, New York, 1963)
5. Fasman, G. D., and Blout, E. R., *J. Am. Chem. Soc.*, **82**, 2262 (1960)
6. Imahori, K., *Biachim. Biophys. Acta*, **37**, 336 (1960)
7. Imahori, K., and Yahara, I., *Biopolymers, Symp.*, **1**, 421 (1964)
8. Bradbury, E. M., Elliott, A., and Hanby, W. E., *J. Mol. Biol.*, **5**, 487 (1962)
9. Ikeda, S., Maeda, H., and Isemura, T., *J. Mol. Biol.*, **10**, 223 (1964)
10. Anufrieva, E. V., Bolotina, I. A., Volchek, B. Z., Illarionova, N. G., Kalikhevich, V. I., Korotkina, O. Z., Mitin, Yu. V., Ptitsyn, O. B., Purkina, A. V., and Eskin, V. E., *Biofizika*, **10**, 918 (1965)
11. Volkenstein, M. V., and Zubkov, V. A., *Biopolymers* (In press)
12. Moffitt, W., Fitts, D. D., and Kirkwood, J. G., *Proc. Natl. Acad. Sci. U. S.*, **43**, 723 (1957)
13. Moffitt, W., and Yang, J. T., *Proc. Natl. Acad. Sci. U. S.*, **42**, 596 (1956)
14. Shechter, E., and Blout, E. R., *Proc. Natl. Acad. Sci. U. S.*, **51**, 695 (1964)
15. Tanford, C., De, P. K., and Taggart, V. G., *J. Am. Chem. Soc.*, **82**, 6028 (1960)
16. Carver, J. P., Shechter, E., and Blout, E. R., *J. Am. Chem. Soc.*, **88**, 2562 (1966)
17. Pysh, E. S., *Proc. Natl. Acad. Sci. U. S.* (In press)
18. Blout, E. R., and Shechter, E., *Biopolymers*, **1**, 565 (1963)
19. Davidson, B., Tooney, N., and Fasman, G. D., *Biochem. Biophys. Res. Commun.*, **23**, 156 (1966)
20. Sarkar, P. K., and Doty, P., *Proc. Natl. Acad. Sci. U. S.*, **55**, 981 (1966)
21. Iizuka, E., and Yang, J. T., *Proc. Natl. Acad. Sci. U. S.*, **55**, 1175 (1966)
22. Townend, R., Kumosinski, T. F., Timasheff, S. N., Fasman, G. D., and Davidson, B., *Biochem. Biophys. Res. Commun.*, **23**, 163 (1966)
23. Carver, J. P., Shechter, E., and Blout, E. R., *J. Am. Chem. Soc.*, **88**, 2550 (1966)
24. Litman, B. J., and Schellman, J. A., *J. Phys. Chem.*, **69**, 978 (1965)
25. Yang, J. T., *Proc. Natl. Acad. Sci. U. S.*, **53**, 438 (1965)
26. Iizuka, E., and Yang, J. T., *Biochemistry*, **4**, 1249 (1965)
27. Moffitt, W., and Moscovitz, A., *J. Chem. Phys.*, **30**, 648 (1959)
28. Moscovitz, A., in *Optical Rotatory Dispersion* (Djerassi, C., Ed., McGraw-Hill, New York, 1960)
29. Moscovitz, A., *Advan. Chem. Phys.*, **4**, 67 (1962)
30. Beychok, S., and Fasman, G. D., *Biochemistry*, **3**, 1675 (1964)
31. Holzwarth, G., Gratzer, W. B., and Doty, P., *J. Am. Chem. Soc.*, **84**, 3194 (1962)
32. Brahm, J., and Spach, G., *Nature*, **200**, 72 (1963)
33. Grosjean, M., and Tari, M., *Compt. Rend.*, **258**, 2034 (1964)
34. Holzwarth, G., and Doty, P., *J. Am. Chem. Soc.*, **87**, 218 (1965)
35. Velluz, L., and Legrand, M., *Angew. Chem. Intern. Ed. Engl.*, **4**, 838 (1965)
36. Timasheff, S. N., Susi, H., Townend, R., Mescanti, L., Gorbunoff, M. J., and Kumosinski, T. F., in *Biopolymer Conformations*, (Ramachandran, G. N., Ed., Academic Press, New York, 1967)
37. Pysh, E. (Private communication)
38. Blout, E. R., Carver, J. P., and Gross, J., *J. Am. Chem. Soc.*, **85**, 644 (1963)
39. Bovey, F. A., and Hood, F. P., *J. Am. Chem. Soc.*, **88**, 2326 (1966)
40. Fasman, G. D., *Biopolymers*, **4**, 509 (1966)
41. Harrington, W. F., and Rao, K. V. (In preparation)
42. Pysh, E. S. (In preparation)
43. Tomimatsu, Y., Vitello, L., and Gaffield, W., *Biopolymers*, **4**, 653 (1966)

44. Yang, J. T., *Proc. Natl. Acad. Sci. U. S.*, **53**, 438 (1965)
45. Tomimatsu, Y., and Gaffield, W., *Biopolymers*, **3**, 509 (1965)
46. Schuster, T. M., *Biopolymers*, **3**, 681 (1965)
47. Herskovits, T. T., Townend, R., and Timasheff, S. N., *J. Am. Chem. Soc.*, **86**, 4445 (1964)
48. Timasheff, S. N., Townend, R., and Mescanti, L., *J. Biol. Chem.*, **241**, 1863 (1966)
49. Fasman, G. D. (In preparation)
50. Troitskii, G. V., *Biofizika*, **10**, 895 (1965)
51. Timasheff, S. N., and Susi, H., *J. Biol. Chem.*, **241**, 249 (1966)
52. Beychok, S., *Proc. Natl. Acad. Sci. U. S.*, **53**, 999 (1965)
53. Mommaerts, W. F. H. M., *J. Mol. Biol.*, **15**, 377 (1966)
54. Townend, R., Timasheff, S. N., and Perlmann, G. E. (In preparation)
55. Timasheff, S. N., and Mescanti, L. (In preparation)
56. Miyazawa, T., *J. Chem. Phys.*, **32**, 1647 (1960)
57. Miyazawa, T., and Blout, E. R., *J. Am. Chem. Soc.*, **83**, 712 (1961)
58. Miyazawa, T., in *Polyamino Acids, Polypeptides and Proteins*, 201 (Stahmann, M. A., Ed., Univ. Wisconsin Press, Madison, Wis., 1962)
59. Krimm, S., *J. Mol. Biol.*, **4**, 528 (1962)
60. Crick, F. H. C., and Rich, A., *Nature*, **176**, 780 (1955)
61. Fraser, R. D. B., *Nature*, **170**, 491 (1952)
62. Simmons, N. S., and Blout, E. R., *Biophys. J.*, **1**, 55 (1960)
63. Tsuboi, M., *J. Polymer Sci.*, **59**, 139, (1962)
64. Tsuboi, M., *Biopolymers, Symp.*, **1**, 527 (1964)
65. Suzuki, S., Iwashita, Y., Shimanouchi, T., and Tsuboi, M., *Biopolymers*, **4**, 337 (1966)
66. Fukushima, K., Ideguchi, Y., and Miyazawa, T., *Bull. Chem. Soc. Japan*, **36**, 1301 (1963)
67. Bamford, C. H., Brown, L., Cant, E. M., Elliott, A., Hanby, W. E., and Malcolm, B. R., *Nature*, **176**, 396 (1955)
68. Astbury, W. T., Dickinson, S., and Bailey, K., *Biochem. J.*, **29**, 2351 (1935)
69. Ambrose, E. J., and Elliott, A., *Proc. Roy. Soc. (London), Ser. A*, **208**, 75 (1951)
70. Bradbury, E. M., Brown, L., Downie, A. R., Elliott, A., Fraser, R. D. B., Hanby, W. E., and McDonald, T. R. R., *J. Mol. Biol.*, **2**, 276 (1960)
71. Yahara, I., and Imahori, K., *J. Am. Chem. Soc.*, **85**, 230 (1963)
72. Yahara, I., Imahori, K., Itaka, Y., and Tsuboi, M., *J. Polymer Sci., Pt. B*, **1**, 47 (1963)
73. Fraser, R. D. B., MacRae, T. P., Stewart, F. H. C., and Suzuki, E., *J. Mol. Biol.*, **11**, 706 (1965)
74. Nozaki, Y., and Tanford, C., *J. Biol. Chem.*, **240**, 3568 (1965)
75. Elliott, A., Ambrose, E. J., and Robinson, C., *Nature*, **166**, 194 (1950)
76. Ambrose, E. J., Bamford, C. H., Elliott, A., and Hanby, W. E., *Nature*, **167**, 264 (1951)
77. Elliott, A., Hanby, W. E., and Malcolm, B. R., *Nature*, **180**, 1340 (1957)
78. Beer, M., Sutherland, G. B. B. M., Tanner, K. N., and Wood, D. L., *Proc. Roy. Soc. (London), Ser. A*, **249**, 147 (1959)
79. Elliott, A., Hanby, W. E., and Malcolm, B. R., *Discussions Faraday Soc.*, **25**, 167 (1958)
80. Hamaguchi, K., *J. Biochem. (Tokyo)*, **56**, 441 (1964)
81. Shigorin, D. N., and Zubov, P. I., *Kolloidn. Zh.*, **21**, 636 (1959)
82. Imahori, K., *Biopolymers*, **1**, 563 (1963)
83. Imahori, K., and Momoi, H., *Arch. Biochem. Biophys.*, **97**, 236 (1962)
84. Timasheff, S. N., Townend, R., and Kumosinski, T. F., *J. Biol. Chem.* (In press)
85. Susi, H., Mescanti, L., and Timasheff, S. N. (In press)
86. Timasheff, S. N., Susi, H., and Mescanti, L. (In press)
87. Low, B. W., *J. Polymer Sci.*, **49**, 153 (1960)
88. Singer, S. J., *Proc. Natl. Acad. Sci. U. S.*, **41**, 1041 (1955)
89. Richards, F. M., *Ann. Rev. Biochem.*, **32**, 269 (1963)
90. Horinishi, H., Hachimori, Y., Kurihara, K., and Shibata, K., *Biochim. Biophys. Acta*, **86**, 477 (1964)
91. Hachimori, Y., Kurihara, K., Horinishi, H., Matsushima, A., and

- Shibata, K., *Biochim. Biophys. Acta*, **105**, 167 (1965)
92. Horinishi, H., Kurihara, K., and Shibata, K., *Arch. Biochem. Biophys.*, **111**, 520 (1965)
93. Horinishi, H., Takenaka, O., and Shibata, K., *Arch. Biochem. Biophys.*, **113**, 371 (1966)
- 93a. Sokolovsky, M., and Vallee, B. L., *Abstr. Am. Chem. Soc., 152nd Meeting, New York, September 1966*, C171
94. Margolis, S., and Langdon, R. G., *J. Biol. Chem.*, **241**, 477 (1966)
95. Tabachnick, M., and Sobotka, H., *J. Biol. Chem.*, **235**, 1051 (1960)
96. Tanford, C., *Advan. Protein Chem.*, **17**, 69 (1962)
97. Weil, L., Buchert, A. R., and Maher, J., *Arch. Biochem. Biophys.*, **40**, 245 (1952)
98. Duke, J. A., Bier, M., and Nord, F. F., *Arch. Biochem. Biophys.*, **40**, 424 (1952)
99. Shaw, E., Mares-Guia, M., and Cohen, W., *Biochemistry*, **4**, 2219 (1965)
100. Schoellmann, G., and Shaw, E., *Biochemistry*, **2**, 252 (1963)
101. Weil, L., James, S., and Buchert, A. R., *Arch. Biochem. Biophys.*, **46**, 266 (1953)
102. Koshland, D. E., Jr., *Ann. N. Y. Acad. Sci.*, **103**, 630 (1963)
103. Whitaker, J. R., and Jandorf, B. J., *J. Biol. Chem.*, **223**, 761 (1956)
104. Theorell, H., and Akeson, A., *J. Am. Chem. Soc.*, **63**, 1812, 1818 (1941)
105. Paléus, S., Ehrenberg, A., and Tuppy, H., *Acta Chem. Scand.*, **9**, 365 (1955)
106. Weil, L., and Seibles, T. S., *Arch. Biochem. Biophys.*, **54**, 368 (1955)
107. Kenkare, U. W., and Richards, F. M., *J. Biol. Chem.*, **241**, 3197 (1966)
108. Steiner, R. F., *Arch. Biochem. Biophys.*, **115**, 257 (1966)
109. Witkop, B., *Advan. Protein Chem.*, **16**, 221 (1961)
110. Funatsu, M., Green, N. M., and Witkop, B., *J. Am. Chem. Soc.*, **86**, 1846 (1964)
111. Shaltiel, S., and Patchornik, A., *J. Am. Chem. Soc.*, **85**, 2799 (1963)
112. Ramachandran, L. K., *J. Sci. Ind. Res. (India)*, **21C**, 111 (1962)
113. Green, N. M., and Witkop, B., *Trans. N. Y. Acad. Sci.*, **26**, 659 (1964)
114. Hachimori, Y., Horinishi, H., Kurihara, K., and Shibata, K., *Biochim. Biophys. Acta*, **93**, 346 (1964)
115. Previero, A., Scoffone, E., Benassi, C. A., and Pajetta, P., *Gazz. Chim. Ital.*, **93**, 849 (1963)
116. Koshland, D. E., Jr., Karkhanis, Y. D., and Latham, H. G., *J. Am. Chem. Soc.*, **86**, 1448 (1964)
117. Horton, H. R., and Koshland, D. E., Jr., *J. Am. Chem. Soc.*, **87**, 1126 (1965)
118. Barman, T. E., and Koshland, D. E., Jr., *Abstr. Am. Chem. Soc., 152nd Meeting, New York, September 1966*, C140
119. Okada, Y., Onoue, K., Nakashima, S., and Yamamura, Y., *J. Biochem. (Tokyo)*, **54**, 477 (1963)
120. Ulmer, D. D., *Biochemistry*, **5**, 1886 (1966)
121. Sun, Y.-K., and Chou, C.-L., *Acta Biochim. Biophys. Sinica*, **2**, 194 (1962)
122. Williams, E. J., *Arch. Biochem. Biophys.*, **15**, 21 (1966)
- 122a. Kendrew, J. C., *Brookhaven Symp. Biol.*, **15**, 216 (1962)
123. Green, N. M., *Biochem. J.*, **89**, 609 (1963)
124. Green, N. M., *Biochim. Biophys. Acta*, **59**, 244 (1962)
125. Green, N. M., *Biochem. J.*, **89**, 599 (1963)
126. Green, N. M., *ibid.*, **90**, 564 (1964)
127. Lokshina, L. A., and Orekhovich, V. N., *Biokhimiya*, **29**, 346 (1964)
128. Blumenfeld, O. O., and Perlmann, G. E., *J. Gen. Physiol.*, **42**, 563 (1959)
129. Arnon, R., and Perlmann, G. E., *J. Biol. Chem.*, **238**, 653 (1963)
130. Previero, A., Coletti, M. A., and Galzigna, L., *Biochem. Biophys. Res. Commun.*, **16**, 195 (1964)
131. Galzigna, L., Previero, A., Reggiani, A., and Coletti, M. A., *Experientia*, **20**, 669 (1964)
132. Williams, E. J., and Laskowski, M., Jr., *J. Biol. Chem.*, **240**, 3580 (1965)
133. Oppenheimer, H. L., Mercouroff, J., and Hess, G. P., *Biochim. Biophys. Acta*, **71**, 78 (1963)
134. Viswanatha, T., and Lawson, W. B., *Arch. Biochem. Biophys.*, **93**, 128 (1961)
135. Dasgupta, B. R., Rothstein, E., and Boroff, D. A., *Anal. Biochem.*, **11**, 555 (1965)

136. Wood, H. N., and Balls, A. K., *J. Biol. Chem.*, **213**, 297 (1955)
137. Williams, E. J., Herskovits, T. T., and Laskowski, M., Jr., *J. Biol. Chem.*, **240**, 3574 (1965)
139. Wootton, J. F., and Hess, G. P., *J. Am. Chem. Soc.*, **84**, 440 (1962)
140. Takahashi, T., Hamaguchi, K., Hayashi, K., Imoto, T., and Funatsu, M., *J. Biochem. (Tokyo)*, **58**, 385 (1965)
141. Hartdegen, F. J., and Rupley, J. A., *Biochim. Biophys. Acta*, **92**, 625 (1964)
142. Hayashi, K., Imoto, T., and Funatsu, M., *J. Biochem. (Tokyo)*, **54**, 381 (1963)
143. Hayashi, K., Imoto, T., and Funatsu, M., *ibid.*, **55**, 516 (1964)
144. Laskowski, M., Jr., *Federation Proc.*, **25**, 20 (1966)
145. Viswanatha, T., Lawson, W. B., and Witkop, B., *Biochim. Biophys. Acta*, **40**, 216 (1960)
146. Foss, J. G., *Biochim. Biophys. Acta*, **47**, 569 (1961)
147. Smillie, L. B., and Kay, C. M., *J. Biol. Chem.*, **236**, 112 (1961)
148. Benmouyal, P., and Trowbridge, C. G., *Arch. Biochem. Biophys.*, **115**, 67 (1966)
149. Steiner, R. F., *Biochemistry*, **5**, 1964 (1966)
150. Spande, T. F., and Witkop, B., *Biochem. Biophys. Res. Commun.*, **21**, 131 (1965)
151. Edelhoch, H., and Steiner, R. F., *J. Biol. Chem.*, **240**, 2877 (1965)
152. Crammer, J. L., and Neuberger, A., *Biochem. J.*, **37**, 302 (1943)
153. Riordan, J. F., Wacker, W. E. C., and Vallee, B. L., *Biochemistry*, **4**, 1758 (1965)
154. Kurihara, K., Horinishi, H., and Shibata, K., *Biochim. Biophys. Acta*, **74**, 678 (1963)
155. Bettelheim, F. R., *J. Am. Chem. Soc.*, **76**, 2838 (1954)
156. Feeney, R. E., and Komatsu, S. K., *Abstr. Am. Chem. Soc., 152nd Meeting, New York, September 1966, C222*
157. Ulmer, D. D., *Biochemistry*, **5**, 1886 (1966)
158. Perlmann, G. E., *J. Biol. Chem.*, **241**, 153 (1966)
159. Kronman, M. J. (Private communication)
160. Gorbunoff, M. J. (In press)
161. Gorbunoff, M. J. (To be published)
162. Tachibana, A., and Murachi, T., *Biochemistry*, **5**, 2756 (1966)
163. Matsushima, A., Inada, Y., and Shibata, K., *Biochim. Biophys. Acta*, **121**, 338 (1966)
164. Inada, Y., Matsushima, A., Kamata, M., and Shibata, K., *Arch. Biochem. Biophys.*, **106**, 326 (1964)
165. Hass, L. F., and Lewis, M. S., *Biochemistry*, **2**, 1368 (1963)
166. Simpson, R. T., Riordan, J. F., and Vallee, B. L., *Biochemistry*, **2**, 616 (1963)
167. Fujioka, H., and Imahori, K., *J. Biochem. (Tokyo)*, **53**, 244 (1963)
168. Perlmann, G. E., *J. Biol. Chem.*, **239**, 3762 (1964)
169. Inada, Y., Kamata, M., Matsushima, A., and Shibata, K., *Biochim. Biophys. Acta*, **81**, 323 (1964)
170. Inada, Y., *J. Biochem. (Tokyo)*, **49**, 217 (1961)
171. Herskovits, T. T., *J. Biol. Chem.*, **240**, 628 (1965)
172. Herskovits, T. T., and Laskowski, M., Jr., *J. Biol. Chem.*, **235**, PC56 (1960)
173. Weil, L., Seibles, T. S., and Herskovits, T. T., *Arch. Biochem. Biophys.*, **111**, 308 (1965)
174. Stellwagen, E., *Biochemistry*, **3**, 919 (1964)
175. Nakatani, M., *J. Biochem. (Tokyo)*, **48**, 633 (1960)
176. Herskovits, T. T., and Laskowski, M., Jr., *J. Biol. Chem.*, **237**, 2481 (1962)
177. Rupley, J. A., *Biochemistry*, **3**, 1648 (1964)
178. Flatmark, T., *Acta Chem. Scand.*, **18**, 1796 (1964)
179. Wu, Y. V., and Scheraga, H. A., *Biochemistry*, **1**, 698 (1962)
180. Friedman, M. E., Goldberger, R., and Scheraga, H. A., *Abstr. Am. Chem. Soc., 152nd Meeting, New York, September 1966, C221*
181. Yasunobu, K. T., and Dandliker, W. B., *J. Biol. Chem.*, **224**, 1065 (1957)
182. Riordan, J. F., Sokolovsky, M., and Vallee, B. L., *J. Am. Chem. Soc.*, **88**, 4104 (1966)
183. Riordan, J. F., Sokolovsky, M., and Vallee, B. L., *Abstr. Am. Chem. Soc., 152nd Meeting, New York, September 1966, C172*
184. Edsall, J. T., in *Aspects of Protein Structure*, 179 (Ramachandran, G. N., Ed., Academic Press, London, 1963)

185. Herskovits, T. T., in *Methods in Enzymology*, 8 (Colowick, S. P., and Kaplan, N. O., Eds., Academic Press, New York, in preparation)
186. Kronman, M. J., and Robbins, F., in *Biological Macromolecules*, 3 (Fasman, G. D., and Timasheff, S. N., Eds., Dekker, New York, in preparation, 1967)
187. Riddiford, L. M., *J. Biol. Chem.*, **241**, 2792 (1966)
188. Riddiford, L. M., and Scheraga, H. A., *Biochemistry*, **1**, 95 (1962)
189. Lowey, S., *J. Biol. Chem.*, **240**, 2421 (1965)
190. Lowey, S., and Kucera, J., in *Biochemistry of Muscle Contraction*, (Gergely, J., Ed., Little, Brown, Boston, 1964)
191. Mizukami, H., (Doctoral thesis, Univ. Illinois, Chicago, Ill., 1962)
192. Herskovits, T. T., and Laskowski, M., Jr., *J. Biol. Chem.*, **237**, 3418 (1962)
193. Townend, R., Weinberger, L., and Timasheff, S. N., *J. Am. Chem. Soc.*, **82**, 3175 (1960)
194. Townend, R., and Herskovits, T. T. (In preparation)
195. Timasheff, S. N., and Townend, R., *J. Am. Chem. Soc.*, **83**, 464 (1961)
196. Kronman, M. J., and Holmes, L. G., *Biochemistry*, **4**, 526 (1965)
197. Donovan, J. W., *Biochemistry*, **3**, 67 (1964)
198. Hass, L. F., and Lewis, M. S., *Biochemistry*, **2**, 1368 (1963)
199. Stellwagen, R. H., Riddiford, L. M., and Edsall, J. T., *Abstr. Am. Chem. Soc., 145th Meeting, New York, 1963*, 66C
200. Rickli, E. E., Ghazanfar, S. A. S., Gibbons, B. H., and Edsall, J. T., *J. Biol. Chem.*, **239**, 1065 (1964)
201. Riddiford, L. M., *J. Biol. Chem.*, **240**, 168 (1965)
202. Riddiford, L. M., *ibid.*, **239**, 1079 (1964)
203. Riddiford, L. M., Stellwagen, R. H., Mehta, S., and Edsall, J. T., *J. Biol. Chem.*, **240**, 3305 (1965)
204. Cross, D. G., and Fisher, H. F., *Biochemistry*, **5**, 880 (1966)
205. Cross, D. G., and Fisher, H. F., *Arch. Biochem. Biophys.*, **110**, 222 (1965)
206. Fisher, H. F., and Cross, D. G., *Biochem. Biophys. Res. Commun.*, **20**, 120 (1965)
207. Mercouroff, J., and Hess, G. P., *Biochem. Biophys. Res. Commun.*, **11**, 283 (1963)
208. Hamaguchi, K., Hayashi, K., Imoto, T., and Funatsu, M., *J. Biochem. (Tokyo)*, **55**, 24 (1964)
209. Hamaguchi, K., and Kurono, A., *J. Biochem. (Tokyo)*, **54**, 111 (1963)
210. Hamaguchi, K., Kurono, A., and Goto, S., *J. Biochem. (Tokyo)*, **54**, 259 (1963)
211. Hamaguchi, K., and Kurono, A., *J. Biochem. (Tokyo)*, **54**, 497 (1963)
212. Hamaguchi, K., *J. Biochem. (Tokyo)*, **55**, 333 (1964)
213. Hamaguchi, K., and Imahori, K., *J. Biochem. (Tokyo)*, **55**, 388 (1964)
214. Kurono, A., and Hamaguchi, K., *J. Biochem. (Tokyo)*, **56**, 432 (1964)
215. Hamaguchi, K., *J. Biochem. (Tokyo)*, **56**, 441 (1964)
216. Simpson, R. T., and Vallee, B. L., *Biochemistry*, **5**, 2531 (1966)
217. Schormuller, M. L., (Doctoral thesis, Purdue Univ., Lafayette, Ind., 1965)